NEW SERIES, NO. 143 (VOL. XVIII) SEPTEMBER, 1923

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Formerly the Quarterly Publication of the American Statistical Association

A WEEKLY INDEX NUMBER OF WHOLESALE PRICES

By Irving Fisher

Scarcely more than a half century ago the periodical publication of index numbers began. The minimum period covered was at first the year, then the quarter, and then the month. Promptness of publication has improved from a delay of several months after the period covered to a few weeks or, in some cases, a few days.

The index number described below, and launched at the beginning of this year, is the first general [1] weekly index number to appear. It is published in the press [2] each Monday and covers the week ending the previous Friday noon. The figures are, in fact, calculated on Friday afternoon. This promptness, which gives news value to the index number, has been made possible by certain short-cut methods which will be noted.[3]

The price quotations, which are for 200 commodities, are taken from *Dun's Weekly Review*.

The formula employed is substantially Laspeyre's Formula (No. 53 in the system of numbering used in *The Making of Index Numbers*, the closest approach to the "ideal," No. 353, which the data available permit). This is an "aggregative" formula representing the fluctuations in value of a fixed budget or an imaginary cargo consisting of specified quantities of the 200 goods.

The quantities, or "weights," of the goods included are based on those given by the War Industries Board in the bulletins prepared under the direction of Professor Wesley C. Mitchell and used in his index number for 1,474 commodities for 1913–18. These quantities are per annum and relate to the year 1917.[4]

[1] Weekly indexes for about a score of foods have for some time been published by Gibson and by the *Times Annalist*.

[2] At present it is taken by 74 newspapers, including one in England, one in Canada, and one in Sweden, with an aggregate circulation of about five million, and by 14 other subscribers. By thus distributing its cost, the expense to each subscriber is nominal.

[3] These methods were very briefly outlined in *The Making of Index Numbers*, p. 346.

[4] The 1919 figures of the U. S. Bureau of Labor Statistics now available were not published when the system for the weekly index was being worked out.

The 200 commodities fall under eight classes. These were so selected as to make their relative degrees of importance (as shown by their total values at base, November, 1922, prices) correspond roughly with those of the same classes in the index numbers of the United States Bureau of Labor Statistics and of the War Industries Board. The relative importance of the eight classes is as follows:

RELATIVE IMPORTANCE OF THE CLASSES

Foods	18
Farm products	28
Cloths and clothing	13
Fuel and light	10
Metals and metal products	18
Building materials	7
Drugs and chemicals	3
Miscellaneous	3
	100

To arrive at these figures it was necessary, or at any rate convenient, in many instances to magnify or reduce some of the quantity figures first obtained as above described. That is, some of the eight classes, or parts of classes, were, before such correction, more fully represented by quotations than others. Thus (as compared with, say, "Fuel and Light," "Building Materials," or most of the other groups), the class "Drugs and Chemicals" is only half as fully represented by quotations as it should be. To increase its relative representation, therefore, the original quantity figure of each commodity in it is multiplied by two, *i. e.*, its *correction coefficient*, or "class weight" as Mitchell would say, is 2. In other words, in view of the paucity of quotations for "Drugs and Chemicals," each actual series of quotations is regarded as representing another as well.

CORRECTION COEFFICIENTS USED

Foods and farm products other than hides and leather	0.9
Hides and leather, cloths and clothing	1.25
Metals and metal products	0.8
Drugs and chemicals	2.0

In thus using correction coefficients I follow the example of Mitchell in his War Industries Board Index Numbers.

But the quantity figures resulting from this correction are not the final figures used except in the case of the 28 most important commodities. For the other 172 commodities, the figure thus obtained is still further modified in order to make it a round number—one of the

following: 1,000, 100, 10, 1, .1. In each case the *nearest* [1] round number is taken.

Such a use of round numbers does not reduce the accuracy of the final result because each of the 172 commodities is unimportant and, in so large a number of small errors, there is an almost complete offsetting. Moreover, as was shown in *The Making of Index Numbers* (p. 448), even doubling or halving a weight (which is a greater alteration than usually occurs in this rounding-off process) has a negligible influence on the result. This was found to be true in the doubling or halving test cited in spite of the fact that only 36 commodities were concerned instead of 200. Furthermore, the rounding-off process, in the present case, is confined to unimportant commodities.

But the justification for round weights was not allowed to rest entirely on such general evidence. Actual tests made by comparing the index number for the 200 commodities thus found with that found by using the exact weights showed that the error in the final index number caused by using round weights was less than one per cent. The *week to week* error from using round weights is, of course, even more negligible.

The *base prices* were for the week ending November 17, 1922, and the *base number* taken was not 100 but 156, that being the number of the United States Bureau of Labor Statistics for November, 1922. In this way the weekly index number may be considered as spliced onto the United States Bureau's monthly number. November was chosen as the splicing point because it was the last for which the Bureau's figure was obtainable at the time my own index was launched; and the week ending November 17, 1922, was chosen because it was the central week of November and therefore the one most nearly representing the month.

In this way, via the index number of the United States Bureau of Labor Statistics, the new index gives the price level relatively to 1913 = 100. But, strictly and independently, of course, it gives the price level relatively only to the mid-November week of 1922 = 156.

Table I gives the list of the 200 commodities, the units in which they are quoted, and their quantities, or the number of these units finally included in the imaginary cargo.

The round weights save an enormous amount of labor not only by making easier multiplications but also by reducing the number of multiplications from 172 to 5. Thus, instead of multiplying by 1,000 in each individual case where 1,000 occurs in the table, we first add

[1] Nearest geometrically, not arithmetically. Thus 40 is nearer to 100 than to 10, being above their mean proportional 32.

together all the prices in this entire group (*i. e.*, the group having 1,000 as the quantity) and then multiply their sum by 1,000, once for all. Similarly, the 100 group requires only one multiplication, and likewise the 10 group, the 1 group, and the 0.1 group. As each such multiplication consists merely in changing the decimal point (and in the case of the 1 group not even that) we may say that the multiplication is practically dispensed with in 172 cases, leaving only the 28 multiplications for the important commodities.

Each of the six groups (the 28-group just mentioned and the five other groups with round weights) is segregated and calculated separately by means of a stencil set over Dun's page of price quotations whereby all outside the group are shut off from view. In this way we obtain the aggregate value of each group and so of the entire "cargo" of 200 commodities.

Thus, for the week ending February 9, 1923, wheat was $1.32 per bushel, which makes the entire 603,000,000 bushels worth $796,000,000; hay, being 360,000,000 hundred weight at $1.15 per hundred weight, was worth $414,000,000; furnace coke, 56,700,000 tons at $7.50 per ton, was worth $425,000,000; coffee, No. 7, Rio, 100,000,000 lbs. at 12½ cents per lb. was worth $12,500,000; these with the other 196 articles make the cargo, in the aggregate, worth, in that week, $19,531,810,000.

The *base aggregate* is 19,217.71 million dollars, the value of the cargo at the prices of mid-November last. The index number for any other date is simply the ratio of the value of the cargo, at said other date, to its base value, multiplied by 156, the base number.

In case any commodity is omitted in any week it is of course also

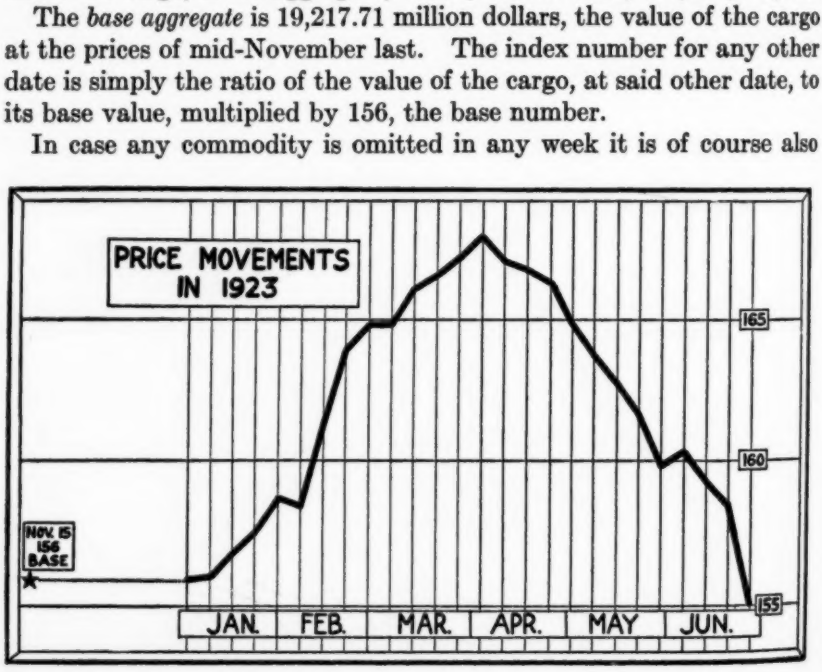


TABLE I

THE 200 COMMODITIES AND THEIR QUANTITIES

TABLE I

THE 200 COMMODITIES AND THEIR QUANTITIES—(000,000 OMITTED)

Commodity	Units in which sold	Number of units (000,000 omitted)
Apples:		
Common	bbl.	52.4
Fancy	"	10.
[1]Beans:		
Marrow, choice	100 lbs.	1.
Pea, choice	" "	1.
Red Kidney, choice	" "	1.
White Kidney, choice	" "	1.
Building material:		
Brick, Hud. R., com.	1,000	10.
Port'd Ct., bulk at mill	bbl.	92.8
Lime, f.o.b. fty	200-lb. bbl.	100.
Burlap, 10½ oz. 40-in	yd.	1,000.
Coffee:		
No. 7 Rio	lb.	100.
Santos No. 4	"	1,000.
Cotton goods:		
Bleach sheetings, st	yd.	1,000.
Medium	"	1,000.
Standard prints	"	1,000.
Brown drills, standard	"	1,000.
Staple ginghams	"	1,000.
Dairy:		
Butter, creamery, extra	lb.	671.
State dairy, com. to fair	"	990.
Cheese, w. n., fresh, spl.	"	100.
Eggs nearby, fancy	doz.	100.
Fresh gathered firsts	"	100.
Dried fruits:		
Apples, evap., choice	lb.	10.
Peaches, Cal., standard	"	100.
Prunes, Cal., 40–50	25-lb. box	100.
Raisins:		
Cal., standard loose mus	lb.	100.
Drugs and chemicals:		
Acetanilid, c.p. bbls	"	1.
Acid, acetic, 28 deg	100 lbs.	1.
Citric, domestic	" "	10.
Nitric, 42°	" "	1,000.
Oxalic	" "	10.
Sulphuric, 60°	" "	100.
Tartaric crystals	lb.	10.
Alcohol, denat. form. 5	gal.	100.
Alum, lump	lb.	1,000.
Arsenic, white	"	100.
Balsam, Peru	"	0.1
Bi-carb'te soda, Am	100 lbs.	10.
Bleaching powder, over 34%	" "	10.
Borax, crystal, in bbl	lb.	100.
Brimstone, crude dom	ton	1.
Calomel, American	lb.	1.
Camphor, domestic	"	10.
Castile soap, pure white	"	100.
Castor oil, No. 1	"	10.
Caustic soda, 76%	100 lbs.	10.
Chlorate potash	lb.	100.
Chloroform	"	10.
Cocaine, hydrochloride	oz.	0.1
Cream tartar, 99%	lb.	0.1
Formaldehyde	"	100.
Glycerine, C. P., in bulk	"	100.
Gum arabic:		
Benzoin, Sumatra	"	1.
Shellac, D. C.	"	10.
Licorice Root	"	10.
Menthol, cases	"	1.
Morphine Sulph., bulk	oz.	1.
Nitrate silver, crystals	"	1.
Nux Vomica, powdered	lb.	10.
Oil Bergamot	"	0.1
Opium, jobbing lots	"	0.1
Quicksilver, 75-lb. flask		0.1
Quinine, 100-oz. tins	oz.	10.
Soda ash, 58% light	100 lbs.	100.
Dyestuffs:		
Gambier	lb.	10.
Indigo, Madras	"	10.
Indigo paste, 20%	"	10.
Fertilizers:		
Bones, ground, steamed 1½% am., 60% bone phosphate, Chicago	ton	0.1
Nitrate soda	100 lbs.	10.
Sulphate, ammonia, domestic, f.o.b. works	" "	10.
Flour:		
Spring Pat	196 lbs.	53.4
Winter, soft straights	" "	53.4
Grain:		
Wheat, No. 2R	bu.	603.
Corn, No. 2, yellow	"	1,260.
Oats, No. 3, white	"	100.
Rye, No. 2	"	100.
Barley, malting	"	100.
Hay, No. 1	100 lbs.	360.
Straw, lg. rye, No. 2	" "	100.
Hemp: Midway, ship	lb.	100.
Hides, Chicago:		
Country No. 1 steers	"	100.
No. 1 buff hides	"	100.
No. 1 extremes	"	100.
No. 1 kip	"	10.
No. 1 calfskin	"	10.
Hops, N. Y., prime '22	"	1.
Leather:		
Union backs, t. r., l. b	"	100.
Scoured oak-backs, No. 1	"	100.
Belting butts, No. 1, t. r., hy	"	100.
[1]Lumber:		
Penn. hemlock, b. price	per M ft.	1.
FAS Pl. wh. oak, 4/4"	" " "	1.
FAS Pl. red gum, 4/4"	" " "	0.1
FAS poplar, 4/4"	" " "	0.1
FAS ash, 4/4"	" " "	0.1
FAS birch, 4/4" (red)	" " "	0.1
FAS chestnut, 4/4"	" " "	0.1
FAS cypress, 4/4" (old grades)	" " "	0.1
FAS H. maple 4/4"	" " "	1.
Adirondack spruce, 2 x 4"	" " "	0.1
No. 1 Com. Y. pine boards, 1 x 4"	" " "	1.
Long leaf yellow pine timbers, 12 x 12"	per M ft.	1.
Douglas fir timbers 12 x 12"	" " "	1.
Metals:		
Pig iron:		
Basic, valley furnace	ton	14.3
Bessemer, Pittsburgh	"	11.9
No. 2, So. Cinc'i	"	4.72
Billets, Bessemer, Pittsburgh	"	8.96
Open-hearth, Phila	"	27.
Wire rods, Pittsburgh	"	1.
O-h. rails, hy., at mill	"	1.
Iron bars, ref., Phila	100 lbs.	10.
Steel bars, Pittsburgh	" "	100.
Tank plates, Pittsburgh	" "	100.
Sheets, black, No. 28, Pittsburgh	" "	34.4
Barb Wire, galvanized, Pittsburgh	" "	10.
Galv. Sheets No. 28, Pittsburgh	" "	10.
Coke Conn'ville, oven,		
Furnace, prompt ship	ton	56.7
Foundry, prompt ship	"	45.5
Aluminum, pig (ton lots)	lb.	100.
Antimony, ordinary	"	10.
Copper, electrolytic	"	1,000.
Lead, N. Y.	"	1,000.
Tin, N. Y.	"	100.
Tin plate, Pittsburgh	100-lb. box	10.
Molasses and syrup:		
Blackstrap	gal.	10.
Ex. fancy	"	10.
Syrup, sugar, medium	"	10.
Naval stores:		
Rosin, "B"	bbl.	10.
Turpentine	gal.	100.
Oils:		
Crude, tks., f.o.b. coast	lb.	100.
China wood, bbls., spot	"	100.
Corn	"	100.
Cottonseed	"	1,000.
Linseed, city, raw	gal.	100.
Palm, Lagos	lb.	100.
Petroleum, cr., at well	bbl.	365.
Kerosene, wagon deliv	gal.	100.
Gas'e auto in gar. st. bbls	"	100.
Soya-bean, tk., coast prompt	lb.	100.
Spot	"	100.
Paints: Litharge, Am	"	100.
Ochre, French	"	10.
Red lead, American	"	100.
White lead in oil	"	10.
White lead dry	"	10.
Whiting Comrcl	100 lbs.	0.1
Zinc, American	lb.	100.
Zinc, F. P. R. S	"	100.
Paper: News roll	100 lbs.	40.1
Book, S. S. & C	lb.	1,000.
Boards, chip	ton	1.
Peas: Scotch, choice	100 lbs.	1.
Platinum	oz.	0.1
Provisions, Chicago		
Beef, live	100 lbs.	122.
Hogs, live	" "	104.
Lard, N. Y. Mid. W	" "	10.
Pork, mess	bbl.	9.
Sheep, live	100 lbs.	10.
Bacon, N. Y., 140s down	" "	1,000.
Hams, N. Y., big, in tcs	" "	1,000.
Tallow, N. Y., sp. loose	" "	100.
Rice:		
Foreign, Saigon No. 1	lb.	1,000.
Rubber:		
Up-river, fine	"	100.
Plan. 1st Latex cr	"	100.
Salt: 280-lb. bbl	bbl.	1.
Salt fish:		
Cod, Grand Banks	100 lbs.	10.
Silk: China, St. Fil 1st	lb.	10.
Japan, Fil., No. 1, Sinshiu	"	38.
Spices:		
Mace	"	1.
Cloves, Zanzibar	"	1.
Nutmegs, 105s–110s	"	1.
Ginger, Cochin	"	1.
Pepper, Singapore, black	"	10.
Sugar: Cent. 96°	100 lbs.	66.2
Fine gran., in bbls	" "	40.2
Tea:		
Formosa, fair	lb.	10.
Fine	"	10.
Japan, low	"	10.
Best	"	1.
Hyson, low	"	10.
Firsts	"	10.
Tobacco, L'ville '21 crop:		
Burley red—Com., sht	"	1,000.
Common	"	100.
Medium	"	100.
Fine	"	10.
Vegetables:		
Cabbage	bbl.	10.
Onions	bag	10.
Wool, Ohio & Pa., fleeces:		
Delaine unwashed	lb.	100.
Half-blood combing	"	100.
Half-blood clothing	"	100.
Wis., Mo. & N. E. fleeces:		
Half-blood	"	10.
Quarter-blood	"	10.
Territory, scoured basis:		
Fine staple choice	"	10.
Half-blood combing	"	10.
Woolen goods:		
Stand. Clay Wor., 16 oz	yd.	50.
Serge, 11-oz	"	50.
Serge, 16-oz	"	50.
Fancy cassimere, 13-oz	"	100.
36-in. all-worsted serge	"	100.
36-in. all-worsted Panama	"	100.
36-in. cotton-warp serge	"	100.

[1] Beans, marrow, included in base aggregate because their omission is very exceptional; and base price taken as $10.00, the price at adjacent weeks.
[2] In carload shipments, f.o.b., New York.

TABLE II

WEEKLY INDEX NUMBERS FOR SIX MONTHS OF 1923

	Index number of prices in per cent of pre-war base, i. e., 1913 = 100 per cent	Purchasing power of the dollar in pre-war cents i. e., 1913 = 100 cents
January—Week ending January 5	155.9	64.1
" " " 12	156.0	64.1
" " " 19	156.8	63.8
" " " 26	157.5	63.5
" " February 2	158.7	63.0
January average	157.0	63.7
February—Week ending February 9	158.4	63.1
" " " 16	161.3	62.0
" " " 23	163.9	61.0
" " March 2	164.8	60.7
February average	162.1	61.7
March—Week ending March 9	164.8	60.7
" " " 16	166.1	60.2
" " " 23	166.6	60.0
" " " 30	167.2	59.8
March Average	166.1	60.2
First Quarterly Average	161.3	62.0
April—Week ending April 6	168.0	59.5
" " " 13	167.1	59.9
" " " 20	166.8	60.0
" " " 27	166.3	60.1
April Average	167.0	59.9
May—Week ending May 4	164.8	60.7
" " " 11	163.7	61.1
" " " 18	162.7	61.5
" " " 25	161.6	61.9
" " June 1	159.8	62.6
May Average	162.5	61.5
June—Week ending June 8	160.3	62.4
" " " 15	159.3	62.8
" " " 22	158.4	63.1
" " " 29	155.0	64.5
June Average	158.3	63.2
Second Quarterly Average	162.6	61.5

omitted from the base aggregate. Thus, white arsenic for the week ending February 2 was missing. Hence its base value, 11.50 million dollars, was subtracted from the base aggregate, making the comparable base 19,206.21, instead of 19,217.71. Consequently, the index number for the week ending February 2 was $\frac{19,531.81}{19,206.21} \times 156 = 158.65$.

The 200 commodities included in this index number are, taken together, about two-thirds as important (as shown by aggregate value) as are the commodities included in the index number of the United States Bureau of Labor Statistics, which are twice as numerous. They

are about a third as important as those of the War Industries Board, which were seven times as numerous. Next to these two, and Dun's monthly index number, the new index number is, I think, the most inclusive of any index number now being published. One or two important commodities like coal are omitted because they are not at present contained in Dun's list. Since Dun's quotations are almost wholly for New York City the index number is practically restricted to New York City.

The purchasing power of the dollar is, of course, readily found from the same comparison above used, inverted.

Table II and the accompanying chart give the index numbers for the first half of 1923. The average for each month is constructed from the figures for the weeks falling wholly or mostly within it. The average is the aggregative average found by summing the values constituting the weekly numerators and the weekly denominators as above explained and dividing the total numerator by the total denominator. The quarterly averages are obtained by the same process.

FIRST AND SECOND LAWS OF ERROR

By Edwin Bidwell Wilson, *Harvard School of Public Health*

In this Journal for March, 1923, Professor W. L. Crum makes an interesting contribution to the use of the median in economic statistics, basing his work as an example upon a series of statistics taken from the *Harvard Review*. He determines three frequency curves to fit the tabulated data, namely, (1) the normal curve of best fit, (2) the Charlier curve of Type A, taking account of skewness and excess but not of the higher terms in the expansion, and (3) the composition of two normal curves. None of these curves fits very well, and the irregularities in the data are such that no particularly good fit could be expected. The general conclusions of Professor Crum are to the effect that the given data and a good many series of data from economic sources probably may be better treated by the median than by the mean, because the median is more stable and because there is close grouping of the majority of the items attended by wide dispersion of the remaining items.

The first two laws of error that were proposed both originated with Laplace. The first law was published in 1774 and stated that the frequency of an error could be expressed as an exponential function of the numerical magnitude of the error, disregarding sign. This may be stated also in the following form: The logarithm of the frequency of an error (all errors being taken without regard to sign) is a linear function of the error. The second law of error was proposed four years later (1778) by Laplace and stated that the frequency of the error is an exponential function of the square of the error; or, stated in another way, that the logarithm of the frequency is a quadratic (parabolic) function of the error. Figure I shows graphically the different characteristics of these two frequency laws.

The second Laplace law is ordinarily called the normal distribution or the Gauss law. It would be just as difficult scientifically to justify the use of the adjective *normal* as it is difficult historically to attribute that law to Gauss, who in spite of his well-known precocity had probably not made this discovery before he was two years old. There are excellent mathematical reasons for the far greater attention which has been paid to the second law than to the first. Both laws are symmetric, but the second, involving the variable x^2, if x be the error, is subject to all the laws of elementary mathematical analysis, whereas the first law, involving as its variable the absolute value of the error x, is

not an analytic function of its variable and therefore presents considerable mathematical difficulty in its manipulation.

No phenomenon is better known perhaps, as a plain matter of fact, than that the frequencies which we actually meet in everyday work in economics, in biometrics, or in vital statistics, very frequently fail to conform at all closely to the so-called normal distribution. The fact

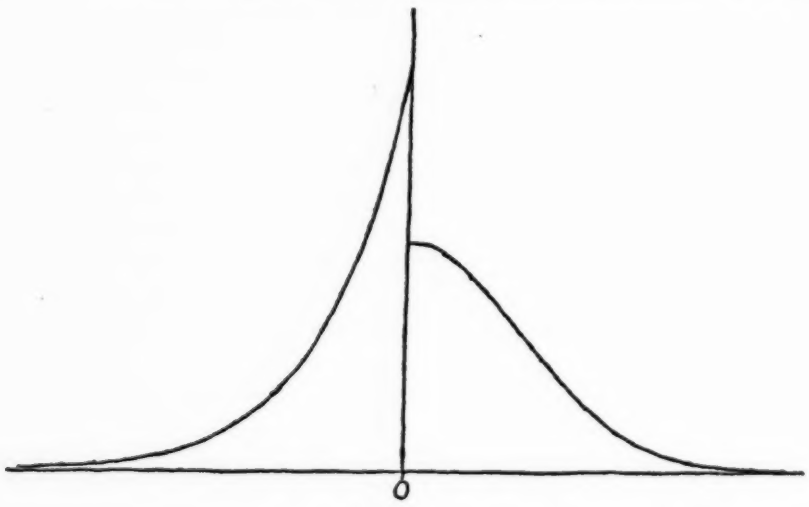


FIGURE I

On the left, Laplace's first frequency curve: $F = \frac{k}{2} e^{-kx}$. On the right, Laplace's second (Gauss's) frequency curve: $F = \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/2\sigma^2}$.

These curves are plotted to a scale which makes the actual probable error the same. Each curve should be reproduced symmetrically on the other side of the central vertical line.

that in the extraordinarily precise measurements of astronomy of position the errors are dispersed about the mean in accordance with the Gauss law is ample justification for the elaborate development of that law in astronomy. The fact that the generalization by Bravais covers very well the observations with respect to the dispersion of shots in artillery and small arms practice is adequate substantiation for a large development of the law of Bravais and the teaching of it in such high-grade institutions as may have to train young men for military or naval affairs. But all this is no justification for attempting to force the law with its various generalizations upon data for which it is not fitted.

One of the first necessities in the discussion of any statistical problem is to study the data themselves. The first attention should be directed toward an examination of the sources of the data to make certain that

the method of compiling the statistical material is well suited to the adequate solution of the problem in hand; and that in particular all considerations which may be germane to the proper selection of material have been applied. Granted that the data have been assembled with care and upon an intelligent system, the next step is to examine the data themselves for the purpose of making some judgment as to the particular statistical treatment which may best fit them. It is by no means safe to rush ahead and apply the second law of Laplace (the law of Gauss) or the various extensions of it which have been developed by the Scandinavian School on the one hand (Gram, Charlier) or the biometric school (Pearson, Yule) on the other. The example upon which Professor Crum has worked is a good case in point, and, largely because the quadratic law for various reasons has such a hold on the statistical community to the unfortunate neglect of the linear law, it is worth while to carry the analysis of this particular set of data a little further and with a somewhat different method of treatment.

The table we need is Professor Crum's Table 2 (page 608):

FREQUENCIES OF DEVIATIONS FROM THE MEDIANS

Deviation	Frequency	Deviation	Frequency	Deviation	Frequency
[1] Over −30	2	−11	6	6	13
−30	1	−10	3	7	8
−29	1	− 9	5	8	6
−28	1	− 8	11	9	5
−24	1	− 7	6	10	2
−23	1	− 6	23	11	4
−22	1	− 5	10	12	3
−21	2	− 4	13	13	1
−20	1	− 3	19	14	2
−19	2	− 2	9	15	1
−18	2	− 1	11	16	1
−17	2	0	28	17	1
−16	2	1	22	18	2
−15	1	2	22	23	1
−14	3	3	13	24	1
−13	6	4	19	28	1
−12	3	5	13	[2] over 30	7

[1] −32, −37.

[2] 34, 35, 35, 41, 41, 42, 45.

The first observation that should be made is that the irregularities of the frequencies are considerable. For example, the deviation of −8 has a frequency of 11, whereas the two adjacent ones have frequencies of 5 and 6; the deviation −6 has a frequency of 23, whereas the adjacent ones have deviations of 6 and 10. In the same way the deviation −3 has a frequency of 19 with the adjacent values 13 and 9; and although we have 28 for zero and 19 for −3, the deviations −1 and −2 have frequencies of 9 and 11. In like manner deviation +3 has a frequency 13, although the frequencies of the adjacent deviations are 22 and 19. Such irregularities as this, if they are found to persist in a

very large series of observations and to occur always at about the same place, indicate of course that there is something permanent in the irregularity; but when the series of observations is small (here only 324) it is simpler to assume that the irregularities are chance fluctuations due to the small number of observations. At any rate, when the irregularities are of the nature that is found here the data are not statistically worth any very careful analysis, and the rather rough methods of treatment used by engineers or economists probably are actually to be preferred on a scientific basis to the more elaborate, and apparently though probably not actually more accurate, analysis which may be had by dissecting the series into the sum of several normal series, or by expanding the series to several terms by the Gram-Charlier method. Such fine distinctions as the difference between the arithmetic mean and the median or the difference between the unadjusted moments and their values as determined by Sheppard adjustments are probably merely a waste of time on tedious arithmetic, unjustified by the data, and may even be worse than a waste of time if we assume from the care and elaborateness of our work that we are really getting ahead any further than if we used rough graphical methods. These comments are made not in criticism of Professor Crum's treatment but rather in support of it, because his main object is the same as ours, namely, to show that the data do not fit the law of the mean and that the use of the median is justified.

When the distribution of any statistical series follows the so-called normal law, the second law of Laplace, there are certain simple relations between the mean deviation, the mean square deviation, the mean cube deviation, etc., which must hold. Let e denote a deviation from a mean. Then

$$nS_1 = e_1 + e_2 + \ldots + e_n, \quad nS_2{}^2 = e_1{}^2 + e_2{}^2 + \ldots + e_n{}^2$$
$$nS_3{}^3 = e_1{}^3 + e_2{}^3 + \ldots e_n{}^3, \quad nS_4{}^4 = e_1{}^4 + e_2{}^4 + \ldots + e_n{}^4$$

Then the ratios of the S's are

$$S_1 : S_2 : S_3 : S_4 = 1.000 : 1.253 : 1.465 : 1.645.$$

And it is quite true, as Bertrand remarks, that "these singular formulas merit so much confidence that a calculator to whom observations have been submitted, and who finds these equalities unsatisfied, may regard it as certain that someone has retouched and altered the immediate results of experiment." These ratios and this statement of Bertrand's apply, however, only to such errors as really obey in the frequency of their distribution the second law of Laplace. It is such a wide experience to persons used to precise measurements that their

errors do follow this law of distribution that Bertrand's dictum is accepted, and properly so, for that kind of material.

Putting the matter another way, one may say, when confronted with data that do not satisfy this continued proportion, that very obviously the data are not distributed in frequency according to the second law. Of course some latitude of departure from the straight proportion must be permitted because the probable errors are of the order of magnitude of $1/\sqrt{n}$ multiplied by these numbers. For example, if the total number of observations is 324, a difference of 5 or even 10 per cent in the computed values from the theoretical values would not be extraordinary.

Now for the data in the table we have approximately

$$S_1 = 7.0,\ S_2 = 10.3,\ S_3 = 13.8,\ S_4 = 17.0.$$

The ratios of these quantities are

$$S_1 : S_2 : S_3 : S_4 = 1 : 1.5 : 2.0 : 2.4.$$

From this it would seem that the test for the normal distribution is very far from being satisfied. The four calculated quantities S_1, S_2, S_3, S_4 are spread out too much, and no reasonable allowance for the operation of probable errors can produce so great a spread.

If the frequency follows the first law of Laplace and varies as e^{-kd} where d is the numerical magnitude of the deviation, a simple application of the integral calculus shows that the ratios of the four S's should be

$$S_1 : S_2 : S_3 : S_4 = 1.000 : 1.414 : 1.817 : 2.213$$

and if we were as accustomed to place our confidence in this law, because of our experience with deviations from a median, as we are to put confidence in the second law, from our experience with the departure of errors from a mean, we should say that unless the four S's calculated from the given data satisfied this continued proportion, except for a reasonable allowance for probable error, we would confidently assume that some one had been tampering with the material. If a comparison be made between the theoretical ratios and ratios for the given data, it appears at once that the distribution in frequency of the data is much nearer, judged by these tests, to Laplace's first law than to the second. It is no longer really unreasonable to maintain that the differences are within the presumptive errors due partly to the scarcity and irregularity of material. There is, however, a little evidence from the progressive nature of the ratios that the observations are more dispersed than they would be even under the first law.

In making the comparison with either law it has been assumed that the distribution is symmetric. This is a fair preliminary assumption because the asymmetry is certainly small. Indeed the departure of

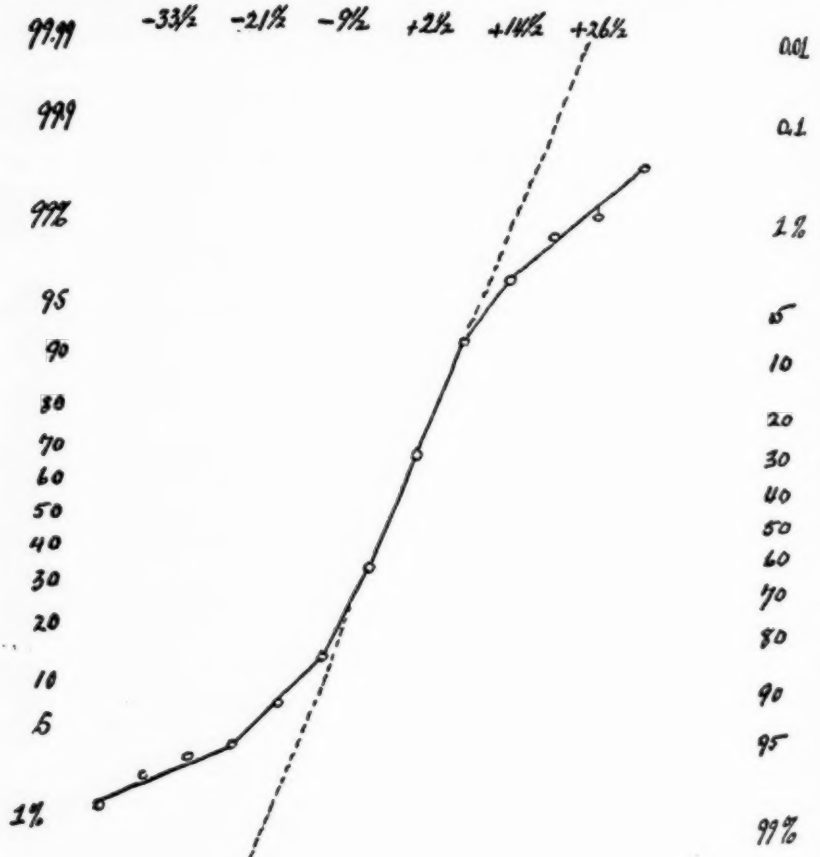


FIGURE II

This chart shows as ordinates on "probability paper" the percentage of deviations which are less than (left scale) or greater than (right scale) a given deviation (plotted as abscissae). If the Gauss law were followed the line would be straight. If the straight line is fitted approximately to the four central points which correspond to the range of greatest frequencies (from $-9\frac{1}{2}$ to $+8\frac{1}{2}$) the probability law would give no deviations in 324 observations greater than +18 or less than −19. If the line be inclined more to give a better fit at the ends, the indicated percentages in the middle range are very far from those actually observed.

the data from symmetry is of about the same order of magnitude measured by its probable error as the departure of the observed S's from the values they should have as calculated from the Gauss law, so that if one can overlook the departures from that law he may perhaps also overlook departures from symmetry.

The numerical calculation of the S's is somewhat annoying, and for the rapidly working statistician who desires to determine with little effort whether the data give any internal evidence of following either of these laws, graphical methods are readily available. We have now probability cross-section paper on which may be plotted the percentage of the deviations which lie within certain limits. If the data follow the second frequency law the points lie approximately upon a straight line. The data have been collected in the table, and Figure II shows how very grossly the points fail to lie upon a straight line.[1] This plot alone would be adequate reason for discontinuing any attempt to represent this material by that law, just as on the numerical side the ratio of the computed S's is justification. The use of arith-log paper enables us to make a similarly simple test for the applicability of the first law.[2] The frequency law may be written as $f=\frac{1}{2}Nk\cdot e^{-kx}$. The total number of observations is N, here 324; the deviation is x; and the number n of deviations beyond a given value x is found by integrating f from x to ∞; $n=\int_x^\infty f dx=\frac{1}{2}Ne^{-kx}$. Hence,

$$\log_{10}n=\log\left(\tfrac{1}{2}N\right)-(k\log_{10}e)x$$

and plots as a straight line on arith-log paper with x as abscissa and n as ordinate. The data in the table have been collected and the results have been plotted in Figure III.

The value of k is the reciprocal $1/\theta$ of the mean deviation θ. If the ratios of $S_1:S_2:S_3:S_4$ had been more nearly in accord with the theoretical values 1:1.4:1.8:2.2 it would be easy to take $\theta=S_1=7.0$. The values of θ calculated respectively from the four S's are 7.0, 7.3, 7.6, and 7.7. The true frequency function is a little larger for large deviations on the positive side and somewhat less for small positive deviations than that obtained from the straight line (see Figure III). However, the departures though apparently systematic are not individually very much greater than might arise from random sampling, and the fit is exceedingly good on the negative side. As we can hardly strive for very great precision with these data we may take

[1] An account of the method of construction and the use of probability paper, with illustrative examples, may be found in G. C. Whipple's *Vital Statistics*, Chap. XII.

[2] The method of the construction and use of arith-log or semi-log paper is described in many elementary texts on graphics.

$\theta = 7.0$ and $k = 1/7$ as a fair representation of the distribution of the data and write $f = 23\ e^{-x/7}$.

The arith-log chart constructed for the Laplacian law, like the probability chart, was based on the total integrated frequency outside a certain limit. This makes the points of very unequal weight, and

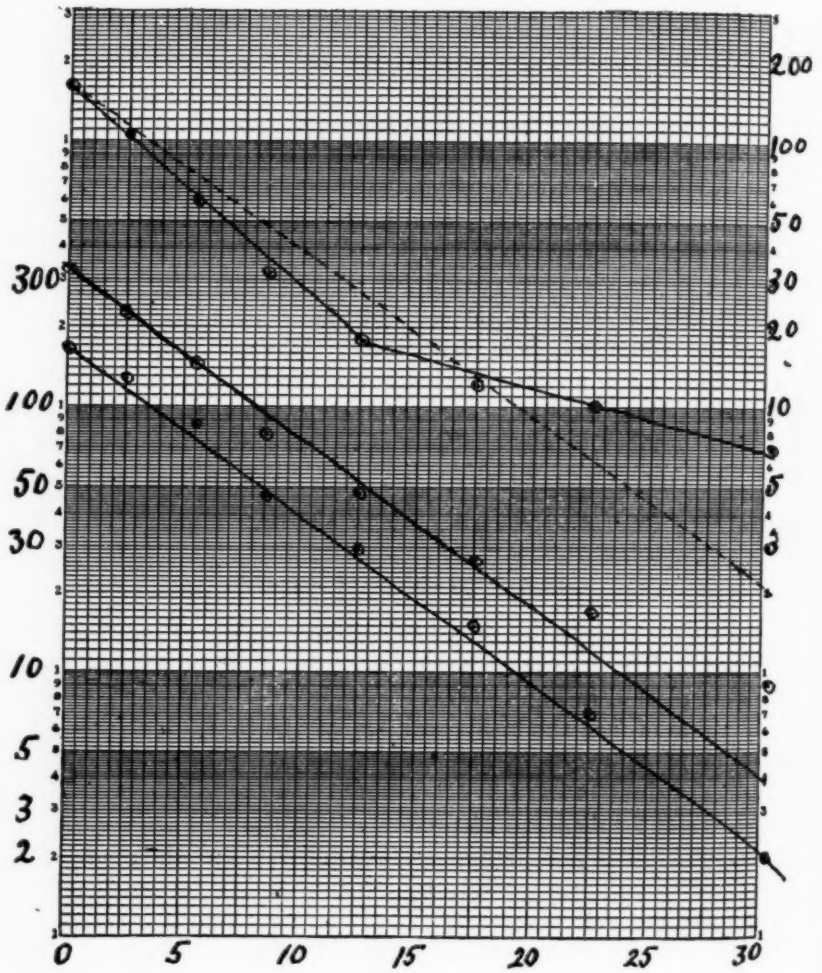


FIGURE III

This chart shows on arith-log paper the number of deviations as ordinates greater than the values given as abscissae. If Laplace's first law holds, the points should lie on a straight line. The lowest set of points and the lowest line are for the negative deviations (left scale), and for them the law holds as well as could be desired. The top line and set of points are for the positive deviations; the fit to the straight dotted line is bad (right scale). The middle line and set of points are for positive and negative deviations taken together (left scale) without regard to sign, and the fit is fair—better than for the probability curve.

each point represents the whole behavior of the curve outside that limit. Hence fluctuations in the curve are somewhat marked. It is possible to construct an arith-log chart from the number of observations between certain equidistant limits. Thus between $x-a$ and $x+a$ the area under the frequency curve is

$$n=\int_{x-a}^{x+a} f dx=\tfrac{1}{2}Nk\int_{x-a}^{x+a} e^{-kx}=dx=\tfrac{1}{2}N(e^{ka}-e^{-ka})e^{-kx}$$

and

$$\log_{10}n=-(k\log_{10}e)x+\log_{10}\tfrac{1}{2}N(e^{ka}-e^{-ka}).$$

If the distribution really follows the Laplacian law, k being constant, the curve obtained by plotting $\log_{10}n$ against x is again straight, and this plot may most easily be made directly on arith-log paper. For the given data, using $a=3$, the tabulated values are:

x	n (obs.)	n (calc.)	x	n (obs.)	n (calc.)
3	109½	(108)	− 3	87½	(90½)
9	33½	(36)	− 9	44	(40)
15	8½	(11½)	−15	16½	(18)
21	3½	(4.0)	−21	8½	(8½)
27	1½	(1.3)	−27	3	(4—)
33	3	(.45)	−33	1½	(1.7)
39	2½	(.15)	−39	1	(0.75)
Total	162	....	Total	162	....

The plot is given in Figure IV. There are two lines which have different slopes and which are practically straight except for the too numerous large positive deviations. With these exceptions the data are well represented by

$$f_1=29.6\,e^{-.183x} \quad \text{for positive deviations}$$
$$f_2=19.4\,e^{-.120x} \quad \text{for negative deviations}$$

with x equal in both cases to the numerical value of the deviations.

The fit of these two frequency curves to the grouped observations, if exception be made of the positive deviations greater than 30, is remarkably good. To have figures comparable to those in Professor Crum's Table IV (page 611) it is merely necessary to divide the numbers in the preceding table by 6. Then

x	obs.	calc.	x	obs.	calc.
3	18.3	18.0	− 3	14.6	15.0
9	5.6	6.0	− 9	7.3	6.7
15	1.4	1.9	−15	2.8	3.0
21	0.6	0.7	−21	1.4	1.4
27	0.25	0.22	−27	0.5	0.65
33	0.50	0.07	−33	0.25	0.28
39	0.41	0.03	−39	0.17	0.12

It is unfair to compare this table with Professor Crum's without calling attention to the fact that here the group interval has been taken as 6 instead of 5, which should give to our table some slight advantage of smoothness of fit; but even with such allowances as might be more than fair, it is difficult for me to believe that the fit here given is not closer than the best he gets (Charlier A type curve). That, however, is not of great importance, for the data are not particularly stable statistically.

The important matters to bear in mind are:

(1) With the exception of the extreme positive deviations which

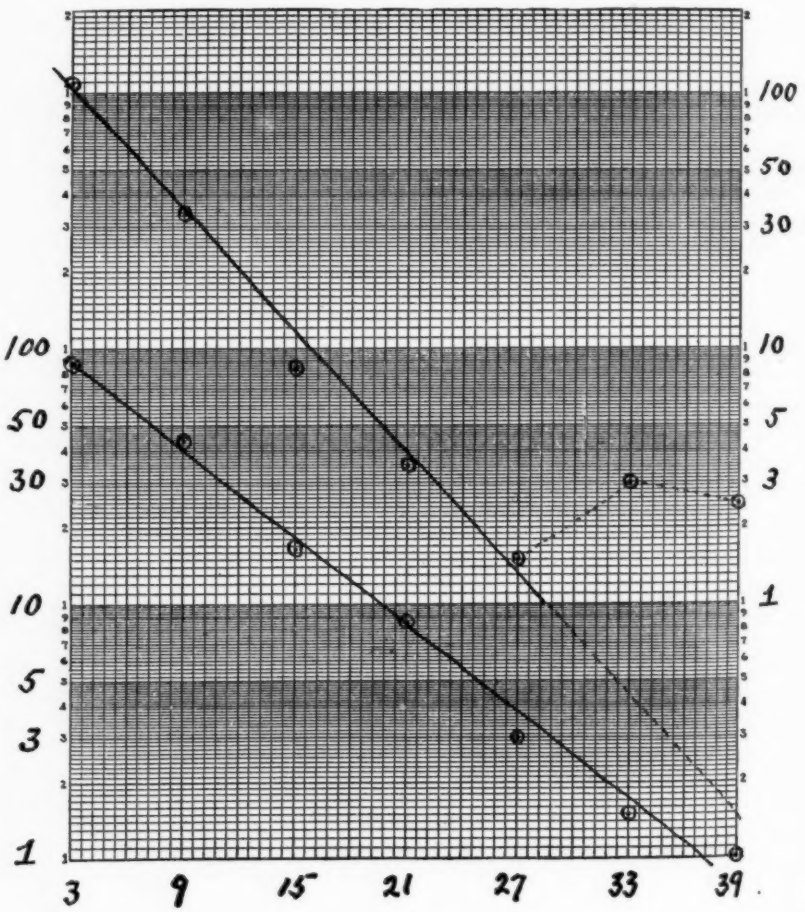


FIGURE IV

This chart shows on arith-log paper in the lower curve the number of negative deviations (left scale) and in the upper curve the number of positive deviations (right scale) in six-unit intervals.

have not been well fitted by any of Professor Crum's three suggestions, these data give internal evidence of following Laplace's first law of error instead of this second law and should be fitted to that law.

(2) By simple graphical means using arith-log paper an extremely good fit may be had in a very few minutes' work (all that is necessary is to plot the grouped data, draw a straight line, and read the graph).

(3) It is not necessary to compute either the mean deviation $\theta = S_1$, or the standard deviation $\sigma = S_2$; for now $\theta = 1/k$ may be read from the chart as equal to 8.3 for negative deviations, and to 5.5 for positive deviations (this in itself taking care of the asymmetry of the distribution) or may be taken as 7.0 for the combined symmetric distributions.

(4) Professor Crum's plea for the use of the median in certain types of statistics is much reinforced by the behavior of these data when discussed in relation to the first law of error.

(5) It is never worth while to spend much time fitting curves of certain types to data until the data have been subjected to careful scrutiny to determine, if may be, what simple laws they do approximately follow.

THE TREND OF ADULT MORTALITY IN THE UNITED STATES[1]

By Gladden W. Baker

There is considerable evidence to show that in the United States the death rate at the higher age periods of life was on the increase for many years previous to 1910. For the registration area as a whole, the death rate was lower in 1900 than in 1890 at every age period up to 60; but after this point in life had been reached the rates were higher in 1900.[2] Dublin has pointed out that the same tendency continued during the following decade.[3] In the original (1900) registration states the male death rate was lower in 1909–11 than in 1900–02 at every year of life up to age 44, and the female rate was reduced up to age 58; but above these ages death rates were higher in the more recent period.[4] When the original registration states are studied individually, nearly all are found to have experienced an increased death rate after middle life in the decade following 1900.[5] Life tables for the cities of Boston, Chicago, New York, and Philadelphia show essentially the same movement.[6] In most of these registration units the increase among males began somewhat earlier in life and was more marked than among females.

Many students of vital statistics have noted the importance of color, nativity, and nationality as factors in mortality, and some have brought these factors to bear on the problem of the increased mortality in adult life. Hoffman found the death rate for the "foreign stock" (that is, foreign-born white persons and their native-born children) higher than the rates for the native stock, especially after age 35; and further distinction showed that the native born of Irish, German, and English maternity suffered conspicuously high death rates.[7] Willcox

[1] This paper contains the results of a dissertation presented to the faculty of the Graduate School in Yale University, 1922, in candidacy for the degree of doctor of philosophy.

[2] Department of the Interior, United States Census Office. Twelfth Census of the United States, Vol. III, p. lxxix.

[3] Dublin, Louis I., "The Trend of American Vitality," *Popular Science Monthly*, 86 (1915): pp. 313–19.

[4] Department of Commerce, United States Bureau of the Census. United States Life Tables: 1890, 1901, 1910, and 1901–10. For the Registration States as a whole, *cf.* pp. 52–63 and charts, p. 244.

[5] Connecticut, Indiana, Massachusetts, Michigan, New Hampshire, New Jersey, New York, Rhode Island, Vermont, and District of Columbia. *Cf.* Institut International de Statistique, *Annuaire International de Statistique*, IV, Mouvement de la population (Amerique), pp. 83–7. *Cf.* also United States Life Tables. Tables, pp. 124–167 and charts, p. 252.

[6] United States Life Tables, pp. 168–99.

[7] Hoffman, Frederick L., "The General Death Rate of Large American Cities," this Journal, 10 (1906–07), pp. 1–75.

has noted that native white persons in New York, 1900 and 1910, had greater vitality at the upper ages than the foreign born.[1] Dublin and Baker found that in New York and Pennsylvania, 1910, the native white stock had lower mortality than the foreign stock.[2] This was true for both sexes and for virtually every age period, being especially marked in adult life. The two branches of the foreign stock agreed much more closely with each other than with the native stock.[3] Considerable variety appeared among the component groups of the foreign born. Austro-Hungarians, Russian Jews, and Italians presented favorable conditions compared with the native stock. The rates for the British, German, and Irish components, however, were not only considerably in excess of the native stock but also much higher than rates in their own native countries. Pulmonary tuberculosis, pneumonia, and the "degenerative diseases" including heart disease, Bright's disease, and cancer were largely responsible for these unfavorable conditions.

This inferior vitality of the foreign born in adult life led to the hypothesis, developed by Dublin and others, that the increasing death rate observed for the total population at the higher ages of life could be explained as the necessary result of the increased numerical importance of the foreign stock at those ages. In his paper, "The Trend of American Vitality," Dublin says:[4]

> This immigration to our registration area must largely determine the adult mortality which these communities experience. If the immigrants are relatively short lived and suffer especially from the diseases of middle life, then we must expect an increased incidence in the mortality rates from these causes in the area where they congregate, and correspondingly a reduction in the expectation of life in the total population. . . . Is it to be wondered at, therefore, that the death rates for large cities and states in the registration area show increases in mortality at the higher age groups? In view of the marked changes that have occurred in the composition and characteristics of our population, it would indeed be surprising if these changes in mortality had not occurred.

In a later study he concludes,[5]

> . . . there is no evidence to support the contention that the true mortality of the country has increased after age 45; . . . the changing constitution of our race stock is responsible for the slight increase in mortality that has occurred.

[1] Willcox, Walter F., New York State Department of Health, Thirty-third Annual Report (1912), "Special Report on Vital Statistics," pp. 240–44.

[2] Dublin, Louis I., and Baker, Gladden W., "The Mortality of Race Stocks in Pennsylvania and New York, 1910," this JOURNAL, 17 (1920–21), pp. 13–44. *Cf.* also Dublin, "Factors in American Mortality," *American Economic Review*, 6 (1916), pp. 523–48.

[3] Pearl, Raymond, "The Vitality of the Peoples of America," *Amer. Jour. of Hygiene*, 1: p. 637. Dr. Pearl finds similar results for six states (Connecticut, Massachusetts, Minnesota, New York, Pennsylvania, and Wisconsin, taken together, 1910).

[4] *Loc. cit.*, pp. 317–18.

[5] "The Increasing Mortality After A[illegible] Forty-Five: Some Causes and Explanations," this JOURNAL, 15: p. 522. *Cf.* also "The Mortality of Foreign Race Stocks," *The Science Monthly*, Jan., 1922, pp. 93–5.

Sir Arthur Newsholme, discussing immigration as a factor in the death rate at the higher ages in the United States, notes that [1]

> . . . nearly one-third of the population of the states of Massachusetts, New York, and New Jersey are foreign born, and of the remaining two-thirds about half have foreign parentage or mixed foreign and native parentage. These foreigners have come in varying proportions from Germany, Ireland, Italy, Russia, and Austria-Hungary, and in many cases they have in earlier life been constantly exposed to the circumstances of malnutrition and insanitation; and, furthermore, in the earlier years of their residence in the United States they possibly have been subjected to excessive strain and privation in insanitary surroundings. It would be surprising if there were not a persistently high or even an increased death rate of persons at higher ages in the registration area.

If it is true that the increase in the death rate of the total population at the higher ages following 1900 is to be explained solely by reference to immigration, an examination of the death rates of the native white

TABLE I

DEATHS PER 1,000 WHITE POPULATION, AMONG: (A) NATIVE BORN OF NATIVE PARENTS; (B) NATIVE BORN OF FOREIGN AND MIXED PARENTS; AND (C) FOREIGN BORN. SEX AND AGE-PERIOD. NEW YORK STATE: 1900, 1910, 1919–20

Sex and age	1919–20			1910			1900		
	(A)	(B)	(C)	(A)	(B)	(C)	(A)	(B)	(C)
Males:									
All ages	13.4	13.5	16.6	15.9	18.7	17.2	17.6	19.5	21.1
10 and over:									
Crude	12.3	11.7	16.7	13.8	13.2	17.5	14.1	12.4	21.2
Standard [a]	12.5	16.4	15.1	13.8	19.5	17.3	13.6	19.1	18.7
Under 10	17.1	17.4	7.9	23.5	31.0	8.4	28.9	34.7	14.7
10–14	2.4	2.4	2.4	2.5	2.2	2.5	2.8	2.9	2.6
15–19	4.0	3.9	4.1	3.6	4.1	4.4	4.4	4.7	4.9
20–24	5.2	5.4	5.7	5.0	6.8	5.2	6.1	8.6	6.9
25–34	6.2	7.8	6.8	5.9	11.6	6.4	6.4	14.0	8.8
35–44	7.6	11.9	9.0	8.2	17.5	11.7	8.6	17.3	13.9
45–54	10.5	18.2	15.9	12.5	23.8	20.7	13.0	20.4	23.5
55–64	22.6	34.7	31.6	27.4	39.4	41.2	24.3	31.2	39.2
65–74	51.9	68.2	65.7	57.7	74.4	74.5	53.3	63.4	71.6
75–84	116.0	156.5	132.5	128.6	163.6	138.5	120.9	160.4	140.2
85 and over	238.9	314.5	245.6	268.9	323.0	272.7	258.0	350.6	270.0
Females:									
All ages	12.5	11.5	16.1	13.9	14.7	16.3	16.1	16.0	19.9
10 and over:									
Crude	12.2	10.4	16.3	12.4	9.7	16.6	13.5	9.8	20.1
Standard [a]	12.3	14.9	14.9	12.4	15.5	16.2	13.2	16.7	18.0
Under 10	13.7	14.2	6.7	19.6	27.0	8.2	24.8	30.0	13.7
10–14	2.1	2.0	2.4	2.6	2.1	2.4	3.0	2.7	2.7
15–19	3.3	3.5	3.8	3.2	3.2	3.2	4.6	4.4	3.7
20–24	5.0	5.5	5.3	4.7	5.2	4.0	6.1	7.5	5.8
25–34	6.2	7.5	6.4	5.1	8.2	5.9	6.5	10.3	8.5
35–44	6.5	8.7	7.4	6.4	10.6	8.9	7.3	11.7	12.1
45–54	9.7	14.2	13.5	10.1	16.4	16.1	11.8	16.6	19.7
55–64	19.5	28.5	28.4	20.1	29.2	34.9	20.6	27.5	36.5
65–74	47.2	61.4	65.6	48.9	62.2	70.3	44.9	56.3	68.4
75–84	106.7	137.5	126.3	113.9	122.9	136.1	110.2	125.3	135.0
85 and over	239.0	261.7	229.2	242.3	324.9	270.5	234.8	390.7	280.2

[a] Standardized for age distribution on basis of native white of native parentage, each sex, in New York, 1910.

[1] "National Changes in Health and Longevity," this JOURNAL, 17: (1920–21), pp. 689–719.

stock and the negro population should reveal no evidences of increase during that decade.

Specific death rates of the three classes of the white population in New York at the last three censuses are presented in Table I [1]. Since the foreign born are but lightly represented at ages 1, 2, and 3 years, where mortality rates are very high, their rate under 10 is not comparable with those of the other two classes. The rates at ages 10 and over have been standardized for age based on the age structure of the native white stock of New York, 1910, which agrees closely with the "standard million" of England and Wales, 1901. The standardized rate of any class indicates what the rate of that class would have been if its age distribution were the same as that of the native stock. It is evident from these standardized rates that the native stock enjoyed much more favorable vitality conditions than the foreign stock, the advantage being more marked among males.

After age 10 the death rates of the three classes are directly comparable. At age periods 10–14 and 15–19, the rates are much the same for the three classes; but beginning with 20–24 and continuing throughout the rest of life, the rates for the native stock are lower than those for the two branches of the foreign stock at each of the census dates. The only exceptions are among foreign-born females at 20–24 in 1900 and 1910, and at 85 and over in 1919–20. The advantage of the native stock becomes greater with advancing age.

The two branches of the foreign stock present a very interesting comparison. At the early age periods their rates are approximately the same, but in middle life native-born children of foreign parents have higher mortality than the foreign born themselves, while at the later ages the situation is reversed. To understand these tendencies consideration must be given to the factor of national extraction. The nationality groups represented in the native born of foreign and mixed parentage class at ages 25–44 in 1910 were the children of German, British, and Irish immigrants.[2] On the other hand, the foreign born class at 25–44 had large proportions of Italians, Russian Jews, and Austro-Hungarians, whose low death rates [3] offset the high mortality among the German, British, and Irish elements and gave the class as a whole a rate lower than that for the native born of foreign and mixed

[1] Population and deaths at "unknown ages" were distributed pro-rata among those at known ages, for each sex and class. Native white decedents of "unknown parentage" were distributed pro-rata among the specified parentage classes, and white decedents of "unknown nativity" were assigned to the specified nativity groups.

[2] A decedent at age 25 in 1910 was evidently born in 1885, and if native born of foreign or mixed parents, his parents belonged to one of the national groups which came to America prior to that time. The year 1883 may be taken as the turning point after which the immigrants from northern and western Europe lost their predominance to the groups which came from southern and eastern Europe.

[3] Dublin and Baker, *loc. cit.*

parentage, which did not yet contain this reducing factor in appreciable proportion. The varying numerical importance from decade to decade of these two contrasting immigrant groups can be traced in the movement of specific mortality rates.

It is clearly evident that the native stock in New York enjoyed much more favorable vitality conditions than the two other classes of white population. But as bearing on the problem of increased mortality at the higher ages, it is more important to know how the mortality of each class compares with its own previous experience than to know how the three classes compare with one another. The changes in death rate among (A) native white of native parents, (B) native white of foreign and mixed parents, and (C) foreign-born white, in New York between 1900 and 1910 may be summarized as follows:

Age period	Males			Females		
	(A)	(B)	(C)	(A)	(B)	(C)
35–44	− 0.4	+ 0.2	−2.2	−1.2	− 1.1	−3.2
45–54	− 0.5	+ 3.4	−2.8	−1.7	− 0.2	−3.6
55–64	+ 3.1	+ 8.2	+2.0	−0.5	+ 1.7	−1.6
65–74	+ 4.4	+11.0	+2.9	+4.0	+ 5.9	+1.9
75–84	+ 7.7	+ 3.2	−1.7	+3.7	− 2.4	+1.1
85 and over	+10.9	−27.6	+2.7	+7.5	−65.8	−9.7

No one class alone was responsible for the increase in death rate of the white population. The mortality record of the native stock during the decade was distinctly less favorable than for the foreign born, the decrease in mortality at 35–44 and 45–54 being less than among the latter group, and the increases thereafter being larger. Although native-born persons of foreign and mixed parents showed the most unfavorable tendencies between 35–44 and 65–74, the native stock suffered greater deterioration at 75–84 and over.

It is true that the composition of the white population at the higher ages changed between 1900 and 1910; in New York the foreign-born element increased in importance in proportion to the two other classes at ages 25–34 and 35–44, while native-born persons of foreign and mixed parentage increased in proportion after 45. It is possible to measure the extent to which these changes would of themselves increase the observed mortality of the total white population at the higher ages. Standardized rates have been constructed for New York which show what the death rate of the white population would have been in 1910 if its nativity and parentage structure had been the same as in 1900.

Age period	Males			Females		
	1900	1910		1900	1910	
		Observed rate	Standard rate		Observed rate	Standard rate
25–34	9.7	7.5	7.9	8.5	6.3	6.4
35–44	13.2	12.2	12.2	10.5	8.7	8.6
45–54	18.8	18.9	18.1	16.0	14.2	13.8
55–64	31.8	35.4	34.9	28.4	28.0	27.7
65–74	62.1	66.8	66.2	56.1	59.7	59.4
75–84	130.0	135.1	133.7	120.7	123.5	124.0

It is evident that these changes in the composition of the white population account for only a small part of the increase in death rates observed at the higher ages between 1900 and 1910; the marked increases shown for males after 55 and for females after 65 are only slightly reduced by this adjustment.

Nor can the increase in death rates of the two branches of the foreign stock be explained as the result of changes in their national elements. During the decade 1900 to 1910, the national groups which predominated among our immigrants before 1883 (Irish, Germans, British) decreased in importance at the upper ages, and their place in the foreign-born class was taken by more recent immigrants who, having come to America in early adulthood, were by 1910 passing into middle life. Remembering the more favorable vitality of the newer immigrant groups, such a displacement should *reduce* rather than increase the death rate of the foreign-born group as a whole. This shifting of national elements, however, would not yet (1910) have affected the foreign born after age period 45–54, nor the native born of foreign and mixed parentage class after 20–24. The increase in rates after these ages occurred in classes whose national composition had remained virtually the same during the decade.

A study of the mortality of the American negro over a period of time is of special interest in relation to the problem of increasing mortality at the higher ages, since this class of the population has remained homogeneous and racially unaffected by immigration. Life tables for negroes in the original (1900) registration states show that in 1909–11 the mortality of negro males, aged 30 to 75, was substantially higher than in 1900–02.[1] Negro death rates are available also for selected registration cities of 10,000 or more population, of which a large part were negroes. In Atlanta, Baltimore, Louisville, Memphis, Mobile, San Antonio, Washington, Atlantic City, Boston, Camden, Chicago,

[1] United States Life Tables, pp. 76–87, and charts, pp. 245, 248.

Cincinnati, Indianapolis, Kansas City (Mo.), and Newark, the negro death rate in 1910 was higher than in 1900 at age period 25–44 or 45–64, or both.[1] The increase in negro death rates in New York, 1900 to 1910, was noteworthy at some of the higher age periods, although it was not so uniform as in the registration states and cities. It appears, therefore, that the deterioration in adult vitality during the decade 1900 to 1910 was not confined to any one class of the population.

Up to this point attention has not been directed to the trend of mortality at the higher ages following 1910. Rates have already been given, however, for the two-year period 1919–20 in New York, and the facts for Pennsylvania, 1910 and 1919–20, are presented in Table II. It is seen that during the decade 1910 to 1920 the adult mortality of each class of the white population was reduced, and in fact every age

TABLE II

DEATHS PER 1,000 WHITE POPULATION, AMONG: (A) NATIVE BORN OF NATIVE PARENTS; (B) NATIVE BORN OF FOREIGN AND MIXED PARENTS; AND (C) FOREIGN BORN. SEX AND AGE-PERIOD. PENNSYLVANIA: 1910, 1919–20

Sex and age	1919–20			1910		
	(A)	(B)	(C)	(A)	(B)	(C)
Males:						
All ages	12.6	14.1	17.4	14.2	21.2	16.4
10 and over:						
Crude	11.4	11.2	17.5	11.5	12.6	16.5
Standard[a]	12.2	15.7	16.0	12.8	18.8	17.5
Under 10	16.8	19.6	7.1	22.6	38.2	11.5
10–14	2.3	2.3	2.2	2.3	2.6	2.6
15–19	3.6	3.8	4.6	3.3	4.1	6.1
20–24	4.5	5.5	7.2	4.9	6.2	7.3
25–34	5.6	7.3	8.3	5.5	10.3	7.5
35–44	7.1	11.5	10.8	7.2	14.7	11.2
45–54	10.3	16.0	17.2	11.4	21.3	19.7
55–64	22.8	32.0	32.1	24.7	35.7	38.3
65–74	53.5	66.3	65.2	57.3	72.2	74.6
75–84	121.3	146.1	127.5	130.2	174.5	142.0
85 and over	249.8	339.9	251.2	253.0	476.5	279.4
Females:						
All ages	12.1	12.4	16.3	12.7	17.3	16.7
10 and over:						
Crude	11.6	10.5	16.5	10.9	10.0	16.9
Standard[a]	12.5	14.9	14.9	12.3	16.3	16.0
Under 10	13.5	16.1	6.1	18.6	32.6	10.8
10–14	2.0	2.0	2.1	2.3	2.2	2.2
15–19	3.4	3.3	3.2	3.3	3.1	3.7
20–24	5.1	6.2	5.6	4.2	5.4	5.0
25–34	6.5	7.7	7.1	5.4	7.7	6.2
35–44	6.8	9.4	8.2	6.3	9.4	8.9
45–54	10.3	13.5	12.6	10.1	14.2	14.3
55–64	20.7	27.7	27.7	20.6	27.0	32.6
65–74	47.2	57.2	63.0	49.5	68.3	67.8
75–84	109.8	135.6	125.9	113.6	152.1	136.9
85 and over	235.1	261.9	241.6	253.8	445.3	276.6

[a] Standardized for age distribution on basis of native white of native parentage, each sex, in New York, 1910.

[1] Department of Commerce, U. S. Bureau of the Census. *Negro Population in the United States*, 1790–1915, pp. 331–35.

period showed a lower death rate in 1919–20 than in 1910, excepting small increases in youth, due to the influenza epidemic of 1919.[1]

It appears from these facts that although the trend of mortality in adult life was upward in the United States prior to 1910, the experience of New York and Pennsylvania gives evidence of a downward movement following 1910.[2] Whether this change in the character of adult mortality is permanent cannot now be judged. It may be significant to note that a similar change has occurred at some time in the registration record of many other areas, including New York City, Massachusetts, and several foreign countries.

A life table is available for the city of New York based on the experience of the two triennia 1879–81 and 1909–11. The mortality of each sex decreased during this 30 year interval through age period 35–39, but at the later ages rates increased.[3] A life table covering the last decade of the 30 year period shows that the death rate of females in New York City was reduced at every age period, and although male rates increased slightly at ages 61 to 64, the increase was very much less than for the three decades combined.[4] These facts indicate that the increases observed in adult life for the 30 year interval occurred before 1900, the changes thereafter being generally favorable.

Specific death rates for Massachusetts from 1876 to 1915, by decades, are given in Table III. It is evident that there were increases at the higher ages in the decade 1886–95 as compared with 1876–85. Among males the increase began with 25–34 and continued through 75–84, while among females the increases began with 35–44. In the decade 1896–1905, the rates at 25–34 and 35–44 receded to points below those from which the rises began. But at 45–54 the reductions were slight in the next two decades, leaving a net increase compared with 1876–85; and at age periods 55–64, 65–74, and 75–84 each decade showed increases over the preceding one.

Some students of adult mortality have assumed that the phenomenon

[1] Ordinarily influenza affects most the extremes of life, and rates are at a minimum between ages 5 and 30. It might be supposed, therefore, that the epidemic of 1918 hastened the death of a large number of old people, thereby reducing the mortality for 1919 and 1920. A number of surveys have shown, however, that this was not the case. The age incidence of influenza was practically the inverse of normal, highest rates occurring at 25–34, with no supernormal mortality in old age. *Cf.* Department of Commerce, U. S. Bureau of the Census. Special Tables of Mortality from Influenza and Pneumonia, 1918, pp. 7–8, and charts pp. 19–24. *Cf.* also Dublin and Frankel, "Influenza Mortality Among Wage Earners and Their Families," *Amer. Jour. of Pub. Health*, 9 (1919): pp. 731–42.

[2] Provisional rates for the expanding registration area, 1910 and 1920, also show decreases at every age period for the decade. *Cf.* U. S. Treasury Department, U. S. Public Health Service, *Public Health Reports*, 37 (March 3, 1922): pp. 487–89.

[3] Department of Health, City of New York. *Monthly Bulletin*, May, 1913.

[4] United States Life Tables, pp. 184–191.

TABLE III

DEATHS PER 1,000 POPULATION, BY AGE AND SEX: MASSACHUSETTS: 1876 TO 1915, BY DECADES[1]

Age	1876–85	1886–95	1896–1905	1906–15
Males:				
All ages[2]	20.1	20.5	18.2	16.4
Standard[3]	21.3	22.5	19.6	17.7
Under 1 year	202.4	247.1	197.2	165.3
1–4	32.8	28.8	19.4	14.6
5–14	6.2	5.2	3.7	3.0
15–24	7.7	7.0	5.8	4.4
25–34	9.4	9.6	7.8	6.5
35–44	10.9	12.0	10.4	9.6
45–54	14.9	16.7	16.4	15.9
55–64	25.8	29.6	31.6	32.8
65–74	54.9	59.4	62.8	66.0
75–84	123.0	125.7	129.5	131.1
85 and over	264.9	257.4	277.8	257.7
Females:				
All ages[2]	19.0	19.1	16.5	14.8
Standard[3]	20.1	20.7	17.5	15.6
Under 1 year	168.7	204.4	154.8	131.6
1–4	31.5	27.3	18.0	13.4
5–14	6.6	5.4	3.6	2.8
15–24	8.7	7.2	5.3	3.8
25–34	10.6	9.3	7.3	5.8
35–44	11.1	11.5	9.4	7.9
45–54	13.5	15.0	14.3	12.8
55–64	22.8	25.8	26.5	26.9
65–74	45.7	52.1	54.8	55.3
75–84	104.2	111.9	114.8	117.2
85 and over	240.9	233.1	255.7	243.2

[1] Institut International de Statistique, *Annuaire International de Statistique:* IV, Mouvement de la population (Amerique), p. 85. (La Haye, 1920.)
[2] Includes unknown age.
[3] International Standard. *Cf. op. cit.*, p. v.

of an increasing death rate at the higher ages was peculiar to the United States. Fisher and Fisk have said:[1]

> This . . . phenomenon, a rising mortality in elderly life, is something almost peculiar to the United States. It is not exhibited in the mortality statistics of the leading European countries. . . . England and Wales, Denmark, Sweden, and Prussia show improved mortality at every age period.

Willcox has expressed the same opinion.[2] Dublin has worked under this assumption in attempting to explain the increased rates in the United States by reference to a condition peculiar to this country, namely, the increased numerical importance at the higher ages of foreign stocks with high death rates.[3] A survey of the trend of mortality in foreign countries does not support these assumptions, but

[1] Fisher, Irving, and Fisk, Eugene. *How to Live: Rules for Healthful Living Based on Modern Hygiene.* New York, 1915, p. 281.
[2] Willcox, Walter F., "The Nature and Significance of the Changes in the Birth and Death Rates in Recent Years," this JOURNAL, 15 (1916–17), pp. 1–15.
[3] Dublin, Louis I., "The Increasing Mortality After Age Forty-Five: Some Causes and Explanations," this JOURNAL, 15 (1916–17), p. 511. *Cf.* also "The Mortality of Foreign Race Stocks," *The Scientific Monthly*, Jan., 1922, pp. 93–103.

shows that several countries have experienced an increasing death rate in adult life at some time during their registration record, many features of which are similar to those noted for Massachusetts.

TABLE IV

DEATHS PER 1,000 POPULATION, BY SEX AND AGE-PERIOD: ENGLAND AND WALES, 1841 TO 1910, BY DECADES[1]

Age	1841–50	1851–60	1861–70	1871–80	1881–90	1891–1900	1901–10
Males:							
All ages[2]	22.5	22.1	22.6	21.8	20.0	19.5	16.6
0—	71.2	72.4	73.5	68.4	61.6	62.7	50.0
5—	9.2	8.5	8.2	6.7	5.3	4.3	3.5
10—	5.1	4.9	4.5	3.7	3.0	2.4	2.1
15—	7.1	6.7	6.2	5.3	4.3	3.8	3.1
20—	9.5	8.8	8.5	7.4	5.7	5.1	4.2
25—	9.9	9.6	9.9	9.3	7.8	6 8	5.6
35—	12.9	12.5	13.5	13.8	12.4	11.5	9.2
45—	18.2	18.0	19.3	20.1	19.4	18.9	16.2
55—	31.8	30.9	33.1	34.9	34.7	34.9	31.8
65—	67.5	65.3	67.1	69.7	70.5	70.4	64.8
75—	148.3	146.7	147.2	150.8	146.6	146.1	137.7
85+	312.3	308.2	315.0	327.4	305.8	286.8	279.2
Females:							
All ages[2]	20.7	20.3	20.2	19.0	17.3	16.7	13.9
0—	61.1	62.7	63.7	58.3	51.9	52.8	41.9
5—	8.9	8.4	7.8	6.2	5.3	4.4	3.6
10—	5.4	5.1	4.5	3.7	3.1	2.6	2.2
15—	7.9	7.4	6.6	5.5	4.4	3.7	2.9
20—	9.1	8.5	8.0	6.8	5.5	4.5	3.5
25—	10.6	9.9	9.7	8.6	7.4	6.1	4.7
35—	12.9	12.2	12.1	11.6	10.6	9.6	7.5
45—	16.0	15.2	15.6	15.6	15.1	14.7	12.5
55—	28.4	27.0	27.9	28.7	28.5	28.4	24.8
65—	61.0	58.7	59.1	61.0	60.4	60.7	53.9
75—	135.9	134.5	134.9	135.4	130.6	130.6	119.8
85+	293.3	288.9	285.1	296.4	270.8	261.4	250.3

[1] Seventy-ninth Annual Report of the Registrar-General of Births, Deaths and Marriages, in England and Wales (1916), pp. 7, 8. (London, 1918).
[2] Standardized to age-constitution of total population as enumerated in 1901.

Specific death rates for England and Wales are shown in Table IV from 1841 to 1910, by decades. It is seen that at ages 25–34 British males suffered a slight increase in the decade 1861–70, but showed improvement in each later decade. At 35–44 the upward trend from 1851 to 1860 culminated in 1871–80, a decade later than the peak at 25–34, and by 1881–90 the rate was reduced to a level slightly lower than that from which the rise began in 1851–60. At each of the next three age periods the increase in death rate also culminated in 1871–80. It is noteworthy that at ages 45–54 and 65–74 it was not until two decades later that the rates were reduced to the level from which the increases began in 1851–60, and at 55–64 this point had not yet been reached by 1901–10. Among females there were no increases at ages 25–34 and 35–44 in 1861–70, though the reduction in mortality was less than in the decade following 1841–50. At 45–54 the rate increased slightly in 1861–70, remained stationary in the following decade, and

in 1881–90 receded to its initial level. At ages 55–64 and 65–74, female rates reached their peak in 1871–80, after which they gradually receded, reaching the level of 1851–60 in 1901–10.[1]

The experience of Scotland from 1856 to 1915 is shown by decades in Table V. In 1866–75 there were increases in death rate for each sex at nearly every age period. This might be accounted for by increased accuracy of registration or by epidemic conditions. Of more significance, therefore, is the experience of the later decades in which mortality at the higher ages increased, in contrast to decreases at the lower ages. For example, the male rate at 55–64 and 65–74 and the female rate at 55–64 increased in every decade through 1896–1905, and though reduced in 1906–15 they were still above the initial level.

TABLE V

DEATHS PER 1,000 POPULATION, BY SEX AND AGE-PERIOD: SCOTLAND, 1856 TO 1915, BY DECADES [a]

Sex and age	1856–65	1866–75	1876–85	1886–95	1896–1905	1906–15
Males:						
All ages [b]	22.6	23.3	20.7	19.4	17.9	16.3
Standard [c]	....	....	21.2	20.5	19.4	17.7
Under 1 year	153.7	159.4	150.1	157.6	158.2	144.0
1–4	38.7	36.8	29.6	27.1	22.7	19.8
5–14	7.4	7.5	6.0	4.6	3.6	3.2
15–24	8.7	8.8	7.3	6.3	5.2	4.3
25–34	10.6	10.9	9.0	8.1	7.2	5.9
35–44	12.6	13.8	12.5	11.7	10.8	8.9
45–54	17.5	20.1	19.4	19.1	18.4	16.3
55–64	29.1	33.0	34.1	34.7	35.5	33.3
65–74	63.4	64.9	65.7	67.4	68.2	65.8
75–84	138.0	142.7	135.3	136.9	135.2	139.0
85 and over	312.9	309.5	305.9	279.9	287.1	282.8
Females:						
All ages [b]	20.6	21.4	19.3	18.5	16.9	15.3
Standard [c]	....	....	19.3	18.9	17.5	16.0
Under 1 year	128.0	130.3	121.1	128.4	126.8	115.1
1–4	38.5	36.0	28.9	26.7	21.7	18.7
5–14	7.4	7.5	6.1	5.0	4.0	3.4
15–24	7.4	8.1	7.1	6.1	5.0	4.1
25–34	9.4	10.0	9.3	8.4	7.1	5.8
35–44	11.4	12.3	11.1	10.9	9.6	8.3
45–54	14.4	16.2	15.2	15.1	15.0	13.2
55–64	24.4	27.0	27.1	28.5	28.7	27.1
65–74	54.5	55.2	54.3	57.5	56.9	53.5
75–84	125.7	129.2	122.9	120.0	117.5	121.6
85 and over	283.1	288.8	274.2	264.8	254.3	253.7

[a] March, L., *Statistique internationale du mouvement de la population*, p. 431, Vol. I. (Paris, 1907) for 1856–65, 1866–75. Thereafter, *cf.* Institut International de Statistique, *Annuaire International de Statistique*, II, Mouvement de la population (Europe), p. 107. (La Haye, 1917.)

[b] Includes unknown age.

[c] International Standard. Cf. *Annuaire*, II, p. viii.

[1] *Cf.* also comparative life tables: Farr's English Life Table No. 3, based on the census enumerations of 1841 and 1851, and upon the deaths registered in the seventeen years 1838–54; Ogle's English Life Table, 1871–80; and Tatham's, 1881–90. Newsholme, Arthur. *The Elements of Vital Statistics*, pp. 299, 307, 314. (London, 3rd Edition, 1899.)

For interesting comment on the English experience made before the reductions in adult mortality had begun to appear, *cf.* Welton, Thos., "On Certain Changes in the English Rates of Mortality," *Jour. of the Stat. Soc.*, 43 (1880), pp. 65–94. Humphreys, Noel, "The Recent Decline in the English Death Rate and Its Effect Upon the Duration of Life," *Jour. of the Stat. Soc.*, 46 (1883), pp. 189–212. Mayo-Smith, Richmond, *Statistics and Sociology*, pp. 178–180, New York, 1910 printing.

The registration record of France, shown in Table VI, also exhibits this contrast. In 1886–95 the male rate was reduced at the younger ages but increased at 35–44 and over, while among females the increases began at 55–64. Although reductions occurred at the higher ages later, the male rate after 55 and the female rate after 75 were still somewhat higher than in 1877–86.

TABLE VI

DEATHS PER 1,000 POPULATION, BY AGE AND SEX: FRANCE: 1877 TO 1913 [1]

Age	1877–86	1886–95	1896–1905	1908–13
Males:				
All ages [2]	23.4	23.3	21.5	19.7
Standard [3]	22.9	22.7	19.8	17.9
Under 1 year	246.6	240.5	185.5	140.4
1–4	29.6	27.2	16.1	12.7
5–14	5.2	4.4	3.6	2.9
15–24	7.3	7.3	6.5	5.6
25–34	9.7	8.7	8.0	7.4
35–44	11.7	11.8	11.0	10.4
45–54	16.7	16.9	16.9	16.6
55–64	29.6	31.3	30.2	30.5
65–74	63.7	68.5	67.6	66.9
75–84	137.4	148.1	163.1	164.7
85 and over	202.1	262.2	350.3	392.4
Females:				
All ages [2]	21.6	21.3	19.3	17.5
Standard [3]	21.1	20.5	17.6	15.4
Under 1 year	205.1	194.5	149.9	114.3
1–4	28.9	26.3	16.6	12.3
5–14	5.7	4.9	4.0	3.2
15–24	6.8	6.5	6.0	5.1
25–34	9.9	8.5	7.5	6.6
35–44	10.4	9.7	8.7	7.8
45–54	13.7	13.3	12.3	11.4
55–64	25.5	26.2	23.6	21.6
65–74	57.6	61.7	57.1	53.4
75–84	128.5	137.1	142.7	140.9
85 and over	210.9	248.8	295.4	329.2

[1] Institut International de Statistique, *Annuaire International de Statistique*: II, Mouvement de la population (Europe), p. 106. (La Haye, 1917.)
[2] Includes unknown age.
[3] International Standard. Cf. *op. cit.*, p. viii.

Hendriks has made the following observations on the mortality experience of Sweden in the century from 1750 to 1850: [1]

The very important tables on the past, compared with the present, rate of mortality of Sweden, are useful in the lesson which they teach us that the rate of mortality, although it has progressively diminished at some ages, has nevertheless increased at certain other ages. . . . Comparing the extremes of this long interval of time shows a diminution of mortality from birth to 36 years of age, but an increase at all ages from 37 years of age to the extremity of life, for males. Comparing the results in the same two periods for female life, there is an apparent diminution, in the one hundred years, of the rate of mortality at each age from birth to 50 years of age, and an increase from 51 years of age to 91.

[1] Hendriks, Frederick, "On the Vital Statistics of Sweden, from 1749 to 1855," *Jour. of the Stat. Soc. of London*, 25 (1862): pp. 118, 125.

In Ireland the increase in death rate during 1876–85, although not confined to the upper ages, was much more marked there than at the younger ages of life.[1] Bulgaria, Serbia, Spain, and Luxemburg have experienced increases in mortality at the higher ages of life, in contrast, generally, with decreases at the lower ages.[2]

SUMMARY

1. In the United States the death rate of the total population at the higher ages of life was on the increase prior to 1910. The inferior vitality of the "foreign stock" in adult life has led some students to the hypothesis that this increase for the total population was due simply to the larger numerical importance of the foreign stock at those ages.

2. A study of the trend of adult mortality in the decade 1900 to 1910 shows this hypothesis to be unfounded:

(a) The increase in adult mortality of New York during the decade was not confined to the foreign white stock. In fact, the mortality experience of the native white stock was distinctly less favorable than that of the foreign born, and at some ages worse even than the first-generation Americans.

(b) The death rate of the white population in New York, 1910, has been "standardized" on the basis of its nativity and parentage composition in 1900. This procedure shows what the mortality rate of the white population would have been in 1910 if no change had occurred in its composition. This standardized rate is only very slightly lower than the observed rate.

(c) In the original registration states the negro population, unaffected by immigration, suffered a large increase in adult mortality between 1900 and 1910.

3. In the decade following 1910 the trend of mortality in New York and Pennsylvania was downward at every age period, including adult life. This suggests that a new phase of adult mortality may have begun in New York about 1910, a period of increasing rates giving way to one of decreasing rates. An analysis of the mortality experience of New York City from 1880 to 1910 and of Massachusetts from 1876 to 1915 indicates a similar change in those registration units.

4. Contrary to the assumption of several writers, the phenomenon of an increasing death rate in adult life is not peculiar to the United States. England and Wales, Scotland, France, Sweden, Ireland, and

[1] March, L. *Statistique internationale du mouvement de la population*, Vol. I, p. 431.

[2] Institut International de Statistique, *Annuaire International de Statistique*, II, Mouvement de la population (Europe), pp. 105–110.

other countries have experienced a similar tendency at some period in their registration history.

5. The adult mortality experiences of England and Wales, Scotland, France, and Massachusetts have several features in common:

(a) The adult death rate of each area appears to have passed through a cycle, a period of increasing rates being followed by one of decreasing rates. In England and Wales the peak occurred in 1871–80, in Scotland during 1866–75, and in France and Massachusetts during 1886–95.

(b) The increase in the death rate began at an earlier age period among males than among females.

(c) At age periods 25–34 and 35–44 the increase occupied only one decade, and in the following decade the rate receded at least as low as the point from which the rise began.

(d) But at the later age periods the increases continued for two or three decades, or even longer; and in some cases the succeeding decreases still left the rates considerably above the initial level.

AN ANALYSIS OF STRIKE STATISTICS, 1881–1921[1]

By Paul H. Douglas, *University of Chicago*

Writing in 1905 in what has ever since remained the classic American text on labor problems, Professor T. S. Adams, after reviewing the strike statistics from 1881 to 1900, declared that "taking one year with another, there is no reason to distrust the plain testimony of the figures that strikes are not increasing as rapidly as the industrial population."[2] From this he went on to say:

> The relative decrease of strikes is undoubtedly due, I think, to improvement in the organization and personnel of the trade-union. . . . The trade-union makes for the regulation not for the suppression of strikes, for their encouragement, in season, for their discouragement out of season, but on the whole its influence is conservative. . . . This calculating regulation has lessened, not only the volume, but even the relative number of strikes. . . . And as our trade unions grow stronger and older it is very probable that the strike will be even more vigorously restricted, because it is the new and poorly organized unions which foment strikes.[3]

These statements of the decline of strikes and their discouragement by trade-unions have had such an influence that it is worth while to examine the basic data upon which these generalizations were made and to determine how well-founded they were at the time they were made and to what degree they have been borne out by subsequent statistics.[4] This attempt to discover the secular trend of strikes is of course different in its aim from such studies as that by Professor Hansen,[5] who has centered his attention upon their cyclical aspects.

The reports of the United States Commissioner of Labor furnish a full record, collected by field agents, of all strikes and lockouts during the period 1881–1905.[6] Unfortunately, our strike statistics were then

[1] I am indebted to my friend and colleague, Professor H. A. Millis, for many helpful criticisms and suggestions.

[2] Adams and Sumner, *Labor Problems*, p. 181.

[3] *Ibid.*, pp. 181–3. Professor Adams seemed to have reached this conclusion because of the fact that "union" strikes decreased in number during the depression years of 1893–97, while "non-union" strikes increased.

[4] Professor Ira B. Cross in an admirable study, "Strike Statistics," this Journal, Vol. XI, pp. 168–94, analyzes the strike statistics from 1881 to 1905. The present study includes the later strike data that have since been made available as well as making several comparisons not covered by Professor Cross. It was impracticable to use the method of the moving average used by Professor Cross because of the nine years' gap between the 1905 and the 1915 statistics.

[5] Alvin H. Hansen, "Cycles of Strikes," *Amer. Economic Rev.*, December, 1921, pp. 615–21.

[6] See the following *Annual Reports of the U. S. Commissioner of Labor:* Third (1888); Tenth (1895); Sixteenth (1901); Twenty-First (1906). It may be asked how these statistics could be collected by field agents for the years 1881–85, since the Bureau of Labor was not created until 1884 and did not begin to function until 1885. The method followed was to make a thorough examination of the files of a wide variety of periodicals published during the years in question and thus to prepare a preliminary

allowed to lapse for an eight-year period. They were not renewed by the Bureau of Labor Statistics until 1914,[1] and have since been carried on to the present time. This later material has been collected by means of press notices and replies to written inquiries rather than by actual field work. It is, therefore, in its very nature not so complete as that for the earlier years; but all its errors lie on the side of minimizing the number of strikes within recent years instead of exaggerating them.[2]

Unfortunately, the figures for 1914 are so unsatisfactory that they cannot be used.[3] Not only is the number of strikes confessedly far below the number that actually occurred, but no statistics at all are given for the number of strikers involved. The statistics for 1915 show a decided improvement, although it is probable that a considerable percentage of the strikes were not recorded. The number of strikers was reported in the case of 752 of the 1,246 strikes. With 1916 the data became more nearly complete. Since then the data have steadily improved in quality. Nevertheless, it should be noted that the number of strikers during the seven years from 1915 to 1921 inclusive was not reported for approximately one-third of the strikes.

Despite all these inadequacies, however, it is possible to secure a tolerably accurate idea of the actual course of strikes during the last forty years. It is evident that the figures given for the recent years are understatements of the actual conditions and hence are most conservative. Table I shows (1) the average number of strikes and strikers by five-year periods since 1881,[4] and (2) the relative number of both strikes and strikers as compared with the average during the first five-year period, 1881–85.

set of data. The country was then districted and field agents were sent out with the preliminary lists. These agents, however, not only gathered information about these specific strikes but they used these and all other available sources to ascertain how many other strikes had occurred during the period. The Bureau declared that "by these means it is believed that the Bureau has secured information relating to nearly every strike, if not every strike, which has occurred in the United States in the period in question." (*Third Annual Report of the U. S. Commissioner of Labor*, p. 11.)

[1] For the strike statistics since 1914, see *The Monthly Labor Review*, July, 1915 (Vol. I), pp. 20–32; April, 1916, pp. 13–22; May, 1922, pp. 181–9.

[2] It may be asked if the use of press notices does not result in an exaggeration of the number of strikers involved. Mr. Ethelbert Stewart, Commissioner of Labor Statistics, replies to this inquiry as follows: "The Bureau of Labor Statistics uses the newspaper report of a strike in most cases simply as an indication of the location of a strike . . . and follows up with a questionnaire to one or both of the parties to the alleged dispute. . . . Thousands of such questionnaires are sent out each year. As to those strikes where information must be accepted from other sources than one or both of the parties to the dispute, the Bureau exercises a most conservative supervision over the published figures of the number of strikes. In other words, the newspapers are used primarily simply to locate strikes and not primarily to obtain information as to the number of people involved." (Letter to the author.)

[3] *Monthly Labor Review*, Vol. I, July, 1915, pp. 20–32.

[4] With the exception of 1914, the seven years 1915–21 are treated as one period.

TABLE I

THE MOVEMENT OF STRIKES IN THE UNITED STATES 1881–1921 [1]

Period	Average number per year of		Relative number (1881–85 = 100) of	
	Strikes	Strikers	Strikes	Strikers
1881–85	498	124,005	100	100
1886–90	1,336	255,863	268	206
1891–95	1,377	278,868	276	225
1896–1900	1,351	281,275	271	227
1901–05	2,793	406,639	561	324
1915–21	3,043	1,547,087 [2]	611	1,248 [2]
		2,395,971 [3]		1,932 [3]
		1,744,502 [4]		1,407 [4]
		1,632,222 [5]		1,316 [5]

[1] Compiled from a table in the *Twenty-First Annual Report of the U. S. Commissioner of Labor* (1906), p. 15; *Monthly Labor Review* (April, 1914), Vol. II, pp. 13–26; (May, 1922), Vol. XIV, pp. 181–6. Only strikes are included; lockouts are not included.
[2] On the basis of Method A.
[3] On the basis of Method B.
[4] On the basis of Method C.
[5] On the basis of Method D.

In attempting to make allowance for the number of strikers involved in those strikes where no such report was made, four distinct and separate methods have been used:

Method A. Only the strikers who were actually reported as having gone out on strike were counted. No allowance was made for the strikers in strikes for which the number of strikers was not reported.

Method B. The average number of strikers involved in such strikes was estimated as being equal to the average number in those strikes for which the number was reported.

Method C. The average number of strikers in such strikes was estimated at 25 per cent of the average number in those strikes for which the reports were made.

Method D. The average number of strikers in these strikes was estimated at only 10 per cent of the number in strikes for which the average was known.

Method A is of course a very real understatement of the number of men actually involved. Method B, on the other hand, is probably an overstatement, since the strikes for which the number of strikers was not reported were probably not so large as those for which full reports were given. Although the estimates of 25 and 10 per cent in Methods C and D respectively are somewhat arbitrary, it is believed that they are on the whole conservative.

These statistics show that during the seven years from 1915 to 1921 inclusive the average number of strikes was slightly over six times as great as during the five years from 1881 to 1885, and between two and three times as great as the average number of strikes during the years

1886–90. Even more notable was the increase in the number of strikers. Counting only those actually reported, with no allowance for workmen involved in the some 8,000 strikes for which the number of strikers was not reported, they totalled nearly twelve and a half times as many as during the first period, while it is probable that in actuality they were somewhere between thirteen and fourteen times as numerous.

It should be noted that there was little increase in the average number of strikes during the fifteen years from 1886 to 1901, and only a 10 per cent increase in the average number of strikers. The period 1901–05, however, witnessed a doubling in the number of strikes and an increase of nearly 50 per cent in the total number of strikers; and since 1905 there has been at least a quadrupling in the number of strikers, although the number of strikes reported has shown a gain of approximately only 20 per cent.

Such statistics do not by themselves, however, give an adequate picture of the real movement of strikes. It is not until they are compared with the increase in the industrial population during these years that their real significance is disclosed. An increase in the total number of strikes or strikers would naturally result from a mere increase in the number "exposed" to strikes, *i. e.*, an increase in the number of industrial wage-earners, and hence would not of itself indicate whether the *rate* of striking was or was not increasing.

The average number of industrial wage-earners during these periods can be computed by following the admirable classification made by Professor Hansen,[1] with an allowance for an even numerical increase during the inter-censal years. The estimated number of industrial wage-earners during these periods is given in Table II.

TABLE II

Period	Estimated average number of industrial wage-earners (in 000's)	Relative number of industrial wage-earners (1881–85 = 100)
1881–85	5,908	100
1886–90	6,905	117
1891–95	8,231	139
1896–1900	9,686	166
1901–05	11,651	197
1916–21	17,030	288

From the foregoing data it is possible to compute the average annual number of strikes and strikers per one million industrial wage-earners together with index numbers of their relative movements, and thus to reduce the strike statistics in Table I to a comparable base.

[1] "Industrial Classes in the United States in 1920," this JOURNAL, Vol. XVII, December, 1922, pp. 503–6.

TABLE III

RELATIVE AVERAGE ANNUAL STRIKE RATES IN RELATION TO NUMBER OF INDUSTRIAL WAGE-EARNERS, BY PERIODS, 1881–1921

Period	Average annual rate per 1,000,000 industrial wage-earners		Relative rates (1881–85 = 100)	
	Strikes	Strikers	Strikes	Strikers
1881–85	84	20,990	100	100
1886–90	193	37,050	230	177
1891–95	167	33,880	199	161
1896–1900	140	29,030	167	138
1901–05	240	34,900	286	166
1915–21	179	90,840 [1]	213	433 [1]
		140,690 [2]		671 [2]
		102,440 [3]		488 [3]
		95,840 [4]		457 [4]

[1] On the basis of Method A.
[2] On the basis of Method B.
[3] On the basis of Method C.
[4] On the basis of Method D.

Some very significant conclusions can be drawn from these statistics. The relative frequency of strikes in relation to the industrial population during the years 1915–21 was more than double what it was for the five years 1881–85. It was, however, approximately 8 per cent less than the rate for the years from 1886 to 1890. The dip in the relative frequency of strikes during the eighteen nineties and their rapid expansion during the first five years of the present century is also clearly evident from these figures.

But more important than the number of strikes is the number of strikers. The proportion of workmen who go out on strike at different times is the best index of whether striking is becoming more or less general in industry. The statistics on this point are even more striking. Whereas during the first half of the eighties approximately only one workman in forty-eight went out on strike each year, during the years 1915–21, at the very least one out of eleven struck, while probably the actual ratio was around one out of ten, or possibly one out of nine. In other words, the proportion of strikers as compared with the number of industrial wage-earners was from four and a third to five times as great during the seven years from 1915 on than it was during the years 1881–85.

Professor Adams's belief in the decline of strikes, based on the statistics up to 1900, is shown to be unfounded if the period 1881–85 is considered as the base, but it is correct if these figures are considered in comparison with the figures for 1886–90. But the increase from 1900 to 1905 which Professor Adams considered as "due to temporary conditions," was apparently not due to temporary causes, for there

was a further increase during the next period for which there is record. It may be objected that the statistics shown for the years 1915–21 cannot be regarded as necessarily indicative of a secular increase in strikes, since the period in question was extraordinary both in the unprecedented price revolution and in the rapid attainment of industrial power by labor. It is lamentable that there are no statistics for the years 1906–14 to determine whether strikes increased appreciably in this country during that period; but the statistics for England, Canada, Australia, and the European Continent indicate a truly remarkable increase in those countries even before the war.[1]

But even the relative number of strikes and strikers in proportion to the industrial population is not sufficient in itself to indicate the relative tendency of strikes to interfere with the course of industry. For duration is just as much a factor in the virulence of strikes as is their scope and frequency. Table IV shows the average annual duration of strikes by periods from 1881 on. The data from 1915 to 1921 are for the 9,849 strikes whose durations were reported. It is possible that the unreported strikes were of shorter duration and that consequently the figures given for the period 1915–21 may somewhat overstate the real situation.

TABLE IV

AVERAGE DURATION OF STRIKES, 1881–1921

Period	Average duration in days	Relative duration (1881–85 = 100)
1881–85	22.7	100
1886–90	23.4	103
1891–95	26.8	118
1896–1900	21.5	95
1901–05	28.2	124
1915–21	28.8	127

It thus appears from the foregoing that strikes are appreciably longer today than they were formerly, and that consequently their intensity is even greater than is indicated by their relative numbers.

What, then, is the truth of what is apparently the other contention in Professor Adams's discussion, namely, that trade-unions tend to decrease the number of strikes? Table V shows the number and the percentage of the strikes which were ordered by labor organizations and of those not ordered by labor organizations, by five-year periods.

[1] For English strike statistics see *The Labour Year Book*, 1916, pp. 189–92, also *Twenty-Sixth Annual Report on Strikes and Lockouts in 1913* (1915 Cd. 7658), *Memoranda relative to Strikes and Lockouts* (1912 Cd. 6081); for Australia, see J. T. Sutcliffe, *A History of Trade Unionism in Australia*, pp. 191–92; for Canada, see the data published from time to time in the *Labour Gazette*.

Unauthorized union strikes are classified among those not ordered by labor organizations.

TABLE V

NUMBER AND PERCENTAGE OF STRIKES ORDERED BY LABOR ORGANIZATIONS AND THOSE NOT SO ORDERED, BY PERIODS FROM 1881 TO 1921[1]

Period	Number of strikes		Percentage of strikes	
	Ordered by labor organizations	Not ordered by labor organizations[2]	Ordered by labor organizations	Not ordered by labor organizations
1881–85	1,311	1,180	53	47
1886–90	4,361	2,316	65	35
1891–95	4,613	2,267	67	33
1896–1900	4,175	2,560	62	38
1901–05	10,893	3,071	78	22
1916–21[3]	12,084	1,902	86	14

[1] Compiled from data in the *Twenty-First Annual Report of the United States Commissioner of Labor*, p. 32; *Monthly Labor Review* (May, 1922), pp. 181–9. Data for 1915 are not available.
[2] "Outlaw" strikes (union strikes not authorized by the organization) are included under the classification of those not ordered by labor organizations.
[3] These figures do not include: (a) 5,894 strikes where the relation to union organization was not reported, and (b) 192 strikes where the employees were not organized at the time of striking but were organized immediately thereafter.

The unions, therefore, from ordering slightly over one-half of the strikes, as in the five years from 1881 to 1885, had come by 1916–21 to ordering a little over seven-eighths.

More important than the relative number of strikes ordered by the unions is the relative number of unionists and non-unionists who went out on strike. One would expect a higher percentage of the strikers to be unionists than of the strikes to be ordered by the unions because of the fact that the union strikes are generally larger than those initiated by non-union workers, and this surmise is borne out by Table VI. Unfortunately, however, the more recent strike statistics do not give the union affiliation of the strikers, and data upon this point exist only for the years 1881–1905.

TABLE VI

NUMBER AND PERCENTAGE OF STRIKERS INVOLVED IN STRIKES ORDERED BY LABOR ORGANIZATIONS AND IN STRIKES NOT SO ORDERED FROM 1881 TO 1905[1]

Period	Average annual number of strikers		Percentage of strikers	
	In strikes ordered by labor organizations	In strikes not ordered by labor organizations	In strikes ordered by labor organizations	In strikes not ordered by labor organizations
1881–85	83,117	40,828	67	33
1886–90	191,203	63,557	75	25
1891–95	223,411	54,029	81	19
1896–1900	219,659	61,613	78	22
1901–05	353,653	52,986	87	13

[1] Computed from data given in the *Twenty-First Annual Report of the U. S. Commissioner of Labor*, p. 42.

It is of course true that the figures of the strikers involved in strikes ordered by the unions include a number of non-unionists. Against this, however, should be set the fact that unauthorized or "outlaw" union strikes and strikers are not included among the strikes "ordered by labor organizations." On the whole it seems probable that the figures given are an approximately accurate measure of the relative proportion of union and non-union strikers.

These figures, however, should be compared with the number and proportion which the unionists formed of the industrial population during each of these periods, if we are to determine the relative liability of trades-unionists and non-trade-unionists to strike. This material is given in Table VII.

TABLE VII

THE ESTIMATED AVERAGE NUMBERS OF TRADES-UNIONISTS BY PERIODS FROM 1881 TO 1921 AND THE PERCENTAGE WHICH THEY FORMED OF THE INDUSTRIAL WAGE-EARNERS[1]

Period	Estimated average number of unionists (in 000's)	Estimated average number of industrial wage-earners (in 000's)	Percentage of unionists to industrial wage-earners
1881–85	300	5,908	5
1886–90	700	6,905	10
1891–95	450	8,231	5
1896–1900	603	9,686	6
1901–05	1,686	11,651	15
1916–21	3,700	17,184	22

[1] For the period 1881–95 the estimates of trade-union membership are based upon data in Commons' *History of Labor in the United States*, Vol. II, pp. 314; 343–4; 351; 396; 423; 473; 482; 494; 501. The membership in the Knights of Labor was as follows:

Year	Membership	Year	Membership	Year	Membership
1881	19,000	1885	105,000	1889	221,000
1882	45,000	1886	710,000	1890	100,000
1883	52,000	1887	501,000	1893	75,000
1884	61,000	1888	260,000		

The membership of the trade-unions outside the Knights of Labor has been estimated by Perlman and Commons at from 200,000 to 225,000 in 1883; at 300,000 in 1885; at something over 250,000 in 1886; and as slightly above this in 1888. In 1893 the membership of those unions affiliated with the A. F. of L. was approximately 275,000 and remained at this figure until 1897. The American Railway Union under Eugene V. Debs claimed 150,000 members in 1894. The membership of the Railway Brotherhoods, the bricklayers, and other unions unaffiliated with the A. F. of L. should be added. On the whole, the estimates made above probably are approximately accurate. Since 1897 the statistics of trade-union membership compiled by Professor George E. Barnett (see *Quar. Jour. of Economics*, Vol. XXX, p. 864, *et seq.*, and *Supplement Amer. Economic Rev.*, Vol. XII, March, 1922, p. 56) have been used. The estimates of the numbers of industrial wage-earners are based upon Professor Hansen's valuable article which has already been cited.

We may now combine the preceding data and compute the relative liability of unionists to strikes as compared with non-unionists.

This relative liability may be computed in terms of either (1) the comparative ratio of strikes ordered and not ordered by labor organiza-

tions or (2) the ratio of union strikers as compared with non-union strikers. The second method is manifestly the more accurate since it deals with actual persons and weights the various strikes by the numbers involved. The first method, however, has the more continuous data and we shall first deal with it and later employ the other method.

On the first basis, this relative liability of unionists to strike is patently the relative number of strikes which a group of unionists (say 100,000 in number) will participate in during a given period in comparison to the number of strikes in which a non-unionist group of equal size will take part. An illustration and a formula will more clearly explain the proper method. If the 10 per cent of the industrial wage-earners who are unionists participate in 50 per cent of all the strikes, while the 90 per cent who are non-unionists take part in only the remaining 50 per cent, then every one per cent of the unionists will be involved in 5 per cent of the strikes, while every one per cent of the non-unionists will take part in only five-ninths of one per cent of the strikes. In proportion to their number, therefore, the unionists will have taken part in nine times as many strikes as have the non-unionists.

The corresponding formula is

$$L = \frac{\frac{S}{U}}{\frac{T}{N}} \times 100.$$

when:

L = Relative liability of unionists to strikes
S = Percentage of strikes ordered by labor organizations
U = Percentage which unionists formed of industrial wage-earners
T = Percentage of strikes originating among non-unionists
N = Percentage which non-unionists formed of industrial wage-earners
100 = Equal liability of unionists to strikes as compared with non-unionists

Table VIII not only gives all the basic data but it shows the relative number of strikes among the unionists in comparison with an equal number of non-unionists.

It is thus apparent that in general the unionists have had approximately twenty times as many strikes relative to their number as have the non-unionists. Once this ratio fell to fifteen, but during the following five years (1891–95) it increased to nearly thirty-nine.

Turning now to the second method, we can secure a more accurate index for the first twenty-five years by substituting the percentage of

TABLE VIII

RELATIVE LIABILITY OF TRADE UNIONISTS TO STRIKES
(100 = EQUAL LIABILITY)

Period	Percentage of strikes ordered by labor organizations	Percentage which unionists formed of industrial wage-earners	Percentage of strikes not ordered by labor organizations	Percentage which non-unionists formed of industrial wage-earners	Relative strike liability of union members (100 = equal liability)
1881–85	53	5	47	95	2,143
1886–90	63	10	37	90	1,532
1891–95	67	5	33	95	3,858
1896–1900	62	6	38	94	2,556
1901–05	78	15	22	85	2,009
1916–21	86	22	14	78	2,178

strikers in union and non-union strikes respectively for the percentage of such *strikes* and designating these percentages by R and V respectively instead of by S and T. The relative liability of unionists to become involved in strikes in comparison with that of an equal number of non-unionists is shown to have been as follows for the period 1881–1905:

Period	Relative liability of unionists to become involved in strikes in comparison with non-unionists (100-equal liability)
1881–85	3,857
1886–90	2,700
1891–95	8,100
1896–1900	5,554
1901–1905	3,792

This shows a much higher ratio even than the former method. In the period (1886–90) in which the disparity in the strike ratio of unionists and non-unionists was least, unionists were apparently involved twenty-seven times as much in strikes as were an equal number of non-unionists, while in one period (1891–95) they were involved approximately eighty times as much.[1] All this, moreover, does not take into account the duration of the strikes, which was undoubtedly much longer in the union than in the non-union strikes because of their better financing, the greater amount of mutual encouragement and the sustained morale which characterized them.

It is lamentable that we have no statistics concerning the number of strikers participating in the union and the non-union strikes for the years 1916–21 inclusive. However, if 87 per cent of all the strikes during this period were ordered by labor organizations, it seems most conservative to estimate that at least 90 per cent of the strikers were involved in those strikes ordered by labor organizations; and if the ratio between the earlier sets of data held true, probably the union strikers would

[1] Adams and Sumner point particularly to the years 1893–97 as proof of the restraining influence of unions upon strikes, yet three of these years are included in the five-year period mentioned above.

comprise an even higher percentage. Using the estimate of 90 per cent, the strike liability of unionists during the period 1916–21 is shown to have been slightly over twenty-five times as great as that of an equal number of non-unionists.

The conclusion seems inescapable that union organization is accompanied by a tremendous increase in the strike rate. This statement should not be interpreted to mean that the unions as institutions incite their members to strike. The non-unionist may want to strike as badly as the unionist, but may be afraid or unable to do so because he has no organization to support him. Unionism, therefore, may merely make possible what is in the minds of all the workers, and may operate, in the language of Bergson, not as an impelling but as a releasing cause.

It may be urged that such statistics do not prove that the unions have caused this higher strike rate among the unionists, since this rate might be due to bad conditions in the trades themselves which might be responsible, indeed, not only for the strikes but for the formation and maintenance of the unions themselves. The force of this argument is doubtful, particularly in view of the generally accepted judgment that the conditions in union plants or crafts are superior to those in non-union establishments. In any event, such an explanation would not be sufficient to explain the great disproportion that exists. It indeed seems indubitable that union organization operates decidedly to increase strikes.

All these facts should not necessarily be interpreted as reasons against unionism. It may well be that industrial peace is not the highest end of industrial relations. It may well be that it is better for the men to organize and to strike more frequently to redress their grievances and to improve their condition than to remain more or less placidly at work and unorganized. This study does not purport to answer such questions. But it is clear that union organization does tend to increase the open manifestations of industrial strife.

The preceding analysis of the strike statistics from 1881 to 1921 may be summarized as follows:

1. The number of strikes in relation to the number of industrial wage-earners was over twice as high during the last seven years as during the period 1881–85. Had all the strikes that actually occurred from 1915 to 1921 inclusive been reported, the increase would have been still more marked. The relative number of these strikes, however, has declined slightly since the period 1886–90, and somewhat more since the period 1901–05.

2. The relative number of strikers in terms of the industrial wage-earning population has shown a truly remarkable increase over that

prevailing in the period 1881–85—being four or five times as great during the years 1915–21 as at that time. This increase is extraordinary when compared with the rates for 1886–90 and for all other periods up to 1905.

3. There has been a distinct tendency for strikes to increase in duration.

4. Strikes are very much more common among union members than among non-unionists. During the years 1916–21 it is probable that unionists tended to become involved in strikes at least twenty-five times as often as non-unionists. From 1881 to 1905 the relative strike liability of unionists ranged from twenty-seven to eighty times that of non-unionists.

HIGHER DEGREES OF COLLEGE PROFESSORS: A STUDY OF SOME FACTORS AFFECTING THE RESEARCH OF GRADUATE STUDENTS WHO BECOME COLLEGE PROFESSORS

By Karl J. Holzinger, *University of Chicago*

In his illuminating work on *American Men of Science*, Cattell has pointed out the grave danger to research due to the relatively late age at which scientific workers receive their higher degrees. He says:

> By our requirements for degrees, by our system of examinations, by our insistence on irrelevant information and ridicule of desirable ignorance and promising mistakes, we crowd on fat when the athlete should be relieved of every superfluous ounce. The doctor's thesis is supposed to be the first productive work; it is completed at the average age of twenty-eight years and is likely to be the working over of the old ideas of an old professor. In the meanwhile the creative instinct has atrophied.[1]

The average man in Cattell's group received his bachelor's degree at twenty-two and his doctor's degree at twenty-eight, so that six years were devoted largely to graduate study. Now this is an impressive fact for the prospective college professor, for it means that the most illustrious men in this profession have taken twice as long to get their higher degrees as the university bulletins would lead one to expect.

The question immediately arises: To what is this prolongation of graduate study due? Is it chiefly because the universities are "crowding on fat," or is it due to a considerable extent to the fact that graduate students marry before they get their doctors' degrees, take masters' degrees, and stop for a time to earn a livelihood? A complete analysis of this problem would involve a study of the personal records of a representative group of students. In the present paper a partial attack on the problem has been made by examining the biographical records of college professors as given in a volume of *Who's Who in America*.

Source and Collection of Data.—This inquiry has been confined to college professors who received a Ph.D. degree at an American University. Professors of law and medicine were excluded because in many cases their graduate work was not comparable with that along academic lines. The following items were then secured for 920 men from the 1916–17 volume of *Who's Who:*

[1] J. McKeen Cattell, *American Men of Science*. New York: The Science Press, 1910, p. 577.

1. Principle subject (now teaching)
2. Age of marriage
3. Age A.B. degree was obtained
4. Age A.M. degree was obtained
5. Age Ph.D. degree was obtained
6. Inter-degree spans

These data were entered on small tickets, and all distributions and tables were made by sorting the cards. The tabulation and the computation throughout are on the basis of one-year intervals with class values at the integral years. The representative character of the data may be seen from the rough classification which follows:

Philosophical group—philosophy, psychology, theology, education	125
Classics	68
Modern languages	53
Social sciences—sociology, anthropology, archeology	119
History	100
Natural science—geology, geography, meteorology	65
Biology	136
Chemistry and physics	96
Mathematics and astronomy	62
English	96
All groups	920

A finer and more accurate subject-grouping was attempted, but was not considered worth while for this study.

Comparison with Cattell's Data.—Cattell's study is confined to illustrious men of science chosen by a rigorous method. The present data are fairly representative of all lines of academic work and were selected by the compilers of *Who's Who* on the basis of "conspicuous achievement" in their respective lines of work.

TABLE I

MEAN AGES AT WHICH A.B. AND PH.D. DEGREES WERE OBTAINED FROM *WHO'S WHO* AND FROM CATTELL'S DATA

Data	A.B. degree		Ph.D. degree	
	Cases	Age	Cases	Age
Who's Who	920	$22.9 \pm .06$	920	$30.2 \pm .12$
Cattell	758	22.2	544	28.4
Difference	...	0.7	...	1.8

Table I presents a comparison of the two sets of data according to the ages at which degrees were received. The probable errors of the means are not indicated in Cattell's results, but assuming that they were even twice as great as ours, the difference in mean ages would still be significant according to the test, $P.E._{a-b} = \sqrt{P.E._a^2 + P.E._b^2}$. It will be noted that the mean age at which the A.B. degree was received is nearly three-quarters of a year greater for the *Who's Who* data, and that the Ph.D. degree was received nearly two years later than in the case of the Cattell group.

Comparison of Subject Groups.—In order to determine whether or not any striking differences occur between the subject groups as defined above, a tabulation was also made dividing the data into those with and without an A.M. degree. The results are given in Table II.

TABLE II

MEDIAN AGES AT WHICH DEGREES WERE OBTAINED AND A.B.–PH.D. SPAN FOR GROUPS WITH AND WITHOUT A.M. DEGREE

Group	Frequency		Age A.B. was obtained		Age Ph.D. was obtained		Span	
	With A.M.	No A.M.	With A.M.	No A.M.	With A.M.	No A.M.	With A.M.	No A.M.
Philosophical	77	48	23.3	22.5	31.9	29.7	7.14	5.93
Classics	36	32	22.4	21.2	30.5	28.5	6.50	6.17
Modern Language	29	24	21.9	21.3	31.0	28.3	7.31	6.83
Social Sciences	58	61	22.3	22.5	29.5	27.1	6.00	4.38
History	50	50	22.4	22.5	30.5	27.8	6.64	5.07
Natural Science	34	31	23.5	22.6	32.2	28.8	6.25	5.75
Biology	79	57	23.3	22.9	29.6	29.1	6.22	5.06
Chemistry	47	49	22.2	23.1	29.0	28.1	6.57	4.81
Mathematics	43	19	22.8	22.7	31.3	28.2	6.29	5.00
English	66	30	21.6	21.5	29.6	26.5	7.14	4.50
All	519	401	22.4	22.3	30.3	28.2	6.68	5.25

The median is used for this table because of the scattering nature of many of the distributions. A close agreement will be noted in the ages at which the A.B. degree was received. The median ages for the Ph.D. degree and corresponding graduate spans are also fairly consistent. Men in the social sciences appear to get their degrees in less time than other groups, but the greatest differences do not appear to be statistically significant; *e. g.*, the difference in graduate span between the philosophical and the social science groups may be written, $(7.14 \pm 0.52) - (6.00 \pm 0.47) = 1.14 \pm 0.70$, which is insignificant because it is not even twice its probable error.

The difference between median ages for groups with and without a master's degree, on the other hand, is clearly significant when the same test is applied. For the "age Ph.D. was obtained" we may write,

$(30.29 \pm 0.21) - (28.15 \pm 0.20) = 2.14 \pm 0.29$. Inasmuch as this difference is about seven times its probable error, the chances are roughly 500,000 to one that it is not due to mere fluctuations in sampling. Obtaining a master's degree unquestionably retards the completion of graduate study by about two years. We shall return to this point in a later paragraph.

Ages at Which Degrees Were Obtained.—The distributions of ages at which all degrees were obtained are presented in Table III. Both medians and means are given because of the definite positive skewness present in all groups. The formula employed in the measurement of skewness was

$$\text{Skewness} = \frac{3\,(\text{mean-median})}{\text{standard deviation}}.$$

TABLE III

DISTRIBUTIONS OF AGES AT WHICH A.B., A.M., AND PH.D. DEGREES WERE OBTAINED BY GROUPS WITH AND WITHOUT A.M. DEGREE

Age	With A.M. degree			Without A.M. degree	
	A.B.	A.M.	Ph.D.	A.B.	Ph.D.
16	..	..	..	1	..
17	1	..	..	3	..
18	8	..	..	7	..
19	22	2	..	10	..
20	65	6	..	41	2
21	69	25	1	73	2
22	102	44	4	81	7
23	66	69	15	66	14
24	62	59	26	42	29
25	35	61	30	33	38
26	27	49	36	15	54
27	17	48	52	13	37
28	16	34	37	5	27
29	8	30	31	1	31
30	9	20	35	1	42
31	6	20	31	1	29
32	1	15	37	2	17
33	1	13	38	3	13
34	2	10	31	1	10
35	2	4	21	..	7
36	..	4	18	..	7
37	..	2	10	1	5
38	..	..	11	..	5
39	..	2	7	1	9
40	..	1	8	..	5
41	..	..	11	..	1
42	..	..	5	..	3
43	..	..	6	..	3
44	..	..	6	..	..
45	..	..	3	..	3
46	..	..	1	..	..
47	..	1	1	..	..
48	..	..	1	..	..
49	..	..	3	..	..
50	..	..	2	..	..
51	..	..	1	..	..
52	..	..	..	..	..
53	..	..	..	..	1
Total	519	519	519	401	401
Mean	23.031	26.133	31.073	22.726	29.030
Median	22.426	25.393	30.286	22.309	28.148
S. D.	2.942	3.823	5.594	2.749	4.820
Skewness	0.617	0.581	0.422	0.455	0.549

Table III reveals a surprising "scatter," especially in the age at which the Ph.D. degree was received. The standard deviation for the bachelor's degree is roughly about half the value for the doctor's degree, a result due to the greater number of retarding factors in the case of the Ph.D.

If the mean be employed as average, the typical man with a master's degree receives his A.B. degree at twenty-three, his A.M. at twenty-six, and his Ph.D. at thirty-one. The typical man without a master's degree receives his Ph.D. at twenty-nine. The corresponding medians are somewhat less on account of the skewness in the distributions. It is obviously necessary to present both means and medians in distributions of this sort.

Inter-degree Spans.—The amount of prolongation in graduate study may be measured by the spans between degrees. Table IV gives distributions of three spans for men with and without a master's degree. If the mean be again used as average, the typical man takes three years

TABLE IV

DISTRIBUTIONS OF SPANS BETWEEN DEGREES FOR GROUPS WITH AND WITHOUT A.M. DEGREE

Span	With A.M. degree			Without A.M.
	A.B.–A.M.	A.M.–Ph.D.	A.B.–Ph.D.	A.B.–Ph.D.
0	..	1	..	..
1	152	50	..	2
2	94	123	14	22
3	144	77	49	76
4	48	61	65	64
5	20	47	68	49
6	15	37	53	39
7	8	22	58	47
8	12	15	29	20
9	10	21	26	19
10	3	15	32	14
11	3	8	21	11
12	4	12	21	7
13	2	6	19	3
14	..	6	13	6
15	..	6	4	4
16	..	2	11	3
17	3	2	4	6
18	..	1	9	4
19	..	..	6	1
20	..	1	4	2
21	..	1	2	..
22	..	3	2	..
23	..	1	1	1
24	..	..	1	..
25	..	1	2	..
26	1	..	1	..
27	..	..	4	1
Total	519	519	519	401
Mean	3.102	4.940	8.042	6.304
Median	2.594	3.639	6.681	5.245
S. D.	2.680	4.020	4.859	3.917
Skewness	0.569	0.971	0.840	0.811

for his master's degree, and five more to obtain his doctorate. If he does not have a master's degree the span is approximately six years. Owing to the skewness of the distributions, the medians differ in some cases by more than a year.

The distribution of the A.B.–A.M. span reveals two crude modes at one and at three years respectively. University catalogs imply that the master's degree requires one year of residence, and our results show that nearly one-third of the men take it in that time. It seems curious, however, that three and not two years mark the next highest frequency. The probable explanation is that if a man fails to get his A.M. in one year he is likely to go out and teach, meanwhile working on his thesis which he brings back to finish the third year.

Percentile Ages and Spans.—The foregoing results are rendered more significant when presented in the form of a percentile table. The fifty percentiles in Table V are of course the medians. Reading across from the thirty percentiles in this table it is apparent that 70 per cent of the men obtained their A.B. degrees at an age later than twenty-one, and their Ph.D. degrees after twenty-seven or twenty-six years of age, according as they had or had not a master's degree.

TABLE V

PERCENTILE AGES AND SPANS FOR GROUPS WITH AND WITHOUT A.M. DEGREE

Percentile	With A.M. degree			Without A.M. degree		
	Age got A.B.	Age got Ph.D.	A.B.–Ph.D. span	Age got A.B.	Age got Ph.D.	A.B.–Ph.D. span
10	19.8	24.7	3.3	20.0	24.0	2.7
30	21.4	27.3	4.9	21.3	26.0	3.8
50	22.4	30.3	6.7	22.3	28.1	5.2
70	24.0	33.2	9.5	23.5	30.4	6.7
90	27.1	38.9	14.4	25.8	35.8	11.3

With even so generous an allowance for graduate study as four or five years, it is evident that nearly 70 per cent took longer and may therefore be considered as retarded.

The ninety percentiles indicate that 10 per cent of the men received their A.B. degrees after twenty-seven and their Ph.D. degrees after the age of thirty-nine, when they also took a master's degree. It is rather surprising that so large a percentage should meet the requirements for *Who's Who* after taking their doctor's degrees at nearly forty.

The Relationship between Age and Span.—The statement has sometimes been made that men who receive their A.B. degrees at an early age tend to complete their graduate work in a relatively short time. The method of studying this problem for the present data was by means

of the product moment coefficient of correlation between the age at which the A.B. degree was obtained and the inter-degree span. Correlation tables with one year intervals were prepared and checked with great care. These tables revealed about the amount of random scatter that one would expect from a handful of sand thrown on such a surface. The coefficients obtained were as follows:

r (age A.B.) (A.B.–Ph.D.) $= +.015 \pm .033$ (N = 401)
r (age A.B.) (A.B.–Ph.D.) $= +.034 \pm .030$ (N = 519)
r (age A.B.) (A.B.–A.M.) $= -.077 \pm .030$ (N = 519)
r (age A.M.) (A.M.–Ph.D.) $= +.017 \pm .030$ (N = 519)

Only in the case of the correlation between age at which the A.B. was received and the A.B.–A.M. span does the coefficient have any significance. This value is negative and about two and one-half times its probable error, so that the odds for significance are about ten to one. We might therefore say that there is a very slight and a not very significant tendency for older men to get their master's degrees in a shorter time. The results as a whole show very clearly that the age at which the A.B. is obtained has no appreciable effect upon the length of the period of graduate study.

Another correlation of interest is that between the graduate spans for the master's and the doctor's degrees. Do men who complete the work for the A.M. degree in a short time also tend to finish their work for the Ph.D. in a brief period? The correlation, r (A.B.–A.M.) (A.M.–Ph.D.) $= +.013 \pm .030$ (N = 519), indicates that no such tendency is present, contrary to popular belief. The two graduate spans are quite independent of one another.

Prediction for Age.—Correlation between the ages at which degrees were received is not in itself of much value because of the spurious result due to the common age interval. Nevertheless, they will furnish the basis for prediction in terms of the regression equation and may therefore be set forth as follows:

r (age A.B.) (age Ph.D.) $= +.583 \pm .023$ (N = 401)
r (age A.B.) (age Ph.D.) $= +.496 \pm .022$ (N = 519)
r (age A.B.) (age Ph.D.) $= +.458 \pm .017$ (N = 920)
r (age A.B.) (age A.M.) $= +.715 \pm .015$ (N = 519)
r (age A.M.) (age Ph.D.) $= +.699 \pm .015$ (N = 519)

In all of the tables the regression was sensibly linear, and we may therefore proceed next to the regression equations themselves:

(1) Age Ph.D. was obtained = 9.33 + .94 (Age A.B. was obtained) ± 3.28 (N = 519), A.M.

(2) Age A.M. was obtained = 4.73 + .93 (Age A.B. was obtained) ± 1.80 (N = 519), A.M.

(3) Age Ph.D. was obtained = 5.83 + 1.02 (Age A.B. was obtained) ± 2.64 (N = 401), No A.M.

(4) Age Ph.D. was obtained = 10.53 + .86 (Age A.B. was obtained) ± 3.22 (N = 920), All.

The interpretation of these equations may be rendered clearer by an example. Suppose a man receives his A.B. at twenty and wishes to determine the most probable age at which he will receive his Ph.D. based on the experience of the above tables. If he intends to stop for a master's degree it will be necessary to use equation (1). Substituting 20 in the right-hand member of this equation we get, Age Ph.D. was obtained = 28.13 ± 3.23. Thus, the most likely age for his Ph.D. will be approximately twenty-eight, but it is an even chance that it will be any age between twenty-five and thirty-one (roughly). The value of the prediction is obviously lessened by the large probable error. In case this man does not stop for a master's degree we use equation (3) giving, Age Ph.D. was obtained = 26.23 ± 2.64, with similar interpretation. Prediction outside of the range of observations given in Table III is extremely unsafe. Regression equations may be used with some success for interpolation, but only very rarely for extrapolation. Such caution must be observed here as in other applications of this method of prediction.

The Marriage Factor.—In examining the 920 cases to determine the age of marriage it was found necessary to discard eleven records because these men either neglected to give the date of marriage or furnished no information at all. From Table VI it is evident that over 90 per cent

TABLE VI

NUMBER AND PER CENT OF MEN MARRIED AND NOT MARRIED WITH AND WITHOUT A.M. DEGREE

Status	With A.M.	Without A.M.
Married	463	372
Not married	46	28
Totals	509	400
Per cent married	90.96	93.00

of the men were married. The difference in the percentage of married men for the groups with and without an A.M. degree may be tested by the formula, $\epsilon_{12}^2 = p_0 q_0 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$. (Yule, p. 269). On applying this test we find, Difference = 93.00 − 90.96 = 2.04 ± 1.83, from which it may

be concluded that the percentages of married men in the two groups are not significantly different.

Tabulations were also made to determine the time of marriage with respect to the year the Ph.D. degree was received. These tables are too bulky to present here, but they are of considerable interest. More men married the same year they took their Ph.D. degrees than at any other time. The mean time of marriage for those without an A.M. was 0.9 years after the Ph.D., while the corresponding time for those with the master's degree was 1.5 years before taking the Ph.D. The mean age at which men with the A.M. married was 29.5, while the corresponding age for those without was 29.9. The difference is not significant, and we may therefore conclude that the typical man in our study married at the age of thirty.

The number of cases reported as "not married" was so small that very reliable conclusions regarding the effect of marriage upon the graduate span cannot be expected. Table VII, however, shows a significant difference for groups with the master's degree. The married men took nearly two years longer to complete their graduate work than those unmarried. Using Pearson's tables it is found that the odds

TABLE VII

A.B.–PH.D. SPANS FOR GROUPS MARRIED AND UNMARRIED WITH AND WITHOUT THE MASTER'S DEGREE

Status	With A.M. degree	Without A.M. degree
Married	8.244±0.157	6.261±0.136
Not married	6.348±0.314	6.607±0.521
Difference	1.896±0.351	−0.336±0.538

against this difference due to random sampling are 3,714 to one. In the case of groups without the master's degree, the difference is insignificant since it is smaller than its probable error.

A final comparison will be made to study the effect of the time of marriage upon the graduate span by dividing the groups into those who married before and those who married after receiving the Ph.D. degree. From the results in Table VIII it is evident that the time of marriage is much more significant than the master's degree in determining the length of the graduate span. Men with the A.M. degree who marry before receiving the Ph.D. take nearly four years longer than those who marry later. For those without the master's degree the corresponding difference is about three years. The differences in the last column show that the master's degree brings about twice as much

TABLE VIII

A.B.–PH.D. SPANS FOR GROUPS MARRIED BEFORE OR AT YEAR OF PH.D. AND FOR THOSE MARRIED AFTERWARD

Time of marriage	With A.M. degree		Without A.M. degree		Difference
	Cases	Span	Cases	Span	
Before or year of Ph.D.	275	9.836 ± 0.215	166	7.861 ± 0.243	1.975 ± 0.324
After Ph.D.	188	5.915 ± 0.166	206	4.971 ± 0.120	0.944 ± 0.205
Difference	463	3.921 ± 0.272	372	2.890 ± 0.271	.....

retardation for those married before taking the Ph.D. as for those married afterward. Late marriage and the absence of the A.M. degree reduce the length of graduate study to five years or one-half of the time spent by the group in early possession of a wife and a master's degree. The typical man in the latter group receives his Ph.D. at about the age of thirty-two, while the man who foregoes early marriage and the intermediate degree may expect to obtain his doctorate at twenty-seven.

Now, it is evident that these gross results do not go far in explaining why retardation should be so great, but they at least point out clearly where it occurs. Early marriage undoubtedly makes it necessary for many men to stop graduate work for a time and earn a livelihood. Furthermore, of those marrying early, 62 per cent obtained a master's degree, while of those marrying late only 48 per cent took this degree. This is probably due to the fact that if a man finds it necessary to teach for a while he is more likely to take the A.M. for the temporary value it has in securing a position.

SUMMARY OF RESULTS

1. The mean age at which men in the *Who's Who* group obtained their Ph.D. degree is nearly two years more than that of Cattell's illustrious scientific men.

2. The average man who took a master's degree received his A.B. at twenty-three, his A.M. at twenty-six, and his Ph.D. at thirty-one.

3. The average man without a master's degree received his A.B. at twenty-three and his Ph.D. at twenty-nine.

4. The typical college professor in our group married at thirty.

5. If he took a master's degree and married while studying, he did not receive his Ph.D. until he was thirty-one. By foregoing the A.M. and early marriage he completed his graduate work at twenty-seven, thereby reducing the period of graduate status from ten to five years.

6. The ages at which A.B. and A.M. degrees were taken are uncor-

related with the length of the graduate span. Furthermore, the A.B.–A.M. and A.M.–Ph.D. spans are themselves uncorrelated, so that prediction for graduate span cannot be made by the method of regression when age or span of the early degrees is given.

7. The ages at which all degrees were obtained show correlations from 0.5 to 0.7 which are due to the common age interval. By the method of regression it is therefore possible to predict the probable age for the higher degrees.

8. The study as a whole shows that early marriage and the taking of a master's degree are the important factors that prolong the period of graduate work. The chief attendant cause is probably the necessity for earning a livelihood during this period. We would not therefore be inclined to blame the university as much as Cattell has done; the fault lies rather in the social conditions under which college professors must study and live.

THE RESEMBLANCE BETWEEN THE ORDINATE OF THE PERIODOGRAM AND THE CORRELATION COEFFICIENT[1]

By W. L. Crum

One of the chief questions to be considered in the analysis of a historical economic series concerns the existence of periodicity in the cyclical fluctuations. A leading method of examining such a series with a view to revealing the presence of a period is the use of the periodogram.[2] The objects of this paper are: to suggest that the periodogram method in its accepted form is not well adapted to the study of economic periodicity, to set forth the close analogy between the periodogram and the correlation methods, and to indicate a revision of the analysis which may be expected to yield a more effective measure of periodicity.

I

In the usual analysis by the periodogram method one constructs a tabular scheme by arranging the ordered items of the given series in successive rows so that each row contains a number of items equal to the trial period to be examined and the number of rows is equal to the integral portion of the quotient of the total number of original items by the trial period. Thus, if the trial period being tested is one of p months and s represents N/p where N is as large a part as possible of the total number of original items, the data will be displayed in a rectangular table containing p columns and s rows. Each column is then totaled and the result divided by s to give the column average. Let the deviation of the average of the ith column from the mean of the N items be designated by $\bar{y}_i$. The ordinate of the periodogram for the trial period p is then R^2, where

$$R^2 = A^2 + B^2 \tag{1}$$

$$\left.\begin{aligned} A &= \frac{2}{p}\sum_{i=0}^{p-1} \bar{y}_i \cos\frac{2\pi}{p} i \\ B &= \frac{2}{p}\sum_{i=0}^{p-1} \bar{y}_i \sin\frac{2\pi}{p} i \end{aligned}\right\} \tag{2}$$

[1] Presented before the American Mathematical Society, April 28, 1923.

[2] For references to several discussions of this method, see: *Rev. Econom. Stat.*, Jan., 1923, p. 19, footnotes.

The method consists in computing R^2 for a range of integral values of p and selecting as the tentative value of the unknown period that value of the trial period p for which R^2 is greatest. Important critical examination is then necessary before this tentative result can be accepted as the true period. It should be remarked that in case the fluctuations are due to several true periods of different lengths, the method will exhibit them all.

It is clear from (1) and (2) that R is the amplitude of the first periodic term in the Fourier expansion fitted to the quantities $\bar{y}_0, \bar{y}_1, \ldots \bar{y}_{p-1}$. If the fluctuations among the $\bar{y}_i$ are cyclical and if the completed cycle has very nearly the form of a sine curve, the first periodic term in the Fourier series will constitute a good fit and R will be truly significant as measuring the amplitude of the actual cycle. If, on the other hand, the cyclical movement among the $\bar{y}_i$ is more complicated, an adequate representation of the fluctuation cannot be obtained from a single sine term: the terms of submultiple period are needed to insure a good fit to the column averages. Although it is clear that we could set up a coefficient similar to R which would measure the extent of the general fluctuation for this more complicated case, it is doubtful if we could entirely justify the use of this new coefficient as a criterion of selection among a range of trial periods. The essential point is that the form of the cyclical fluctuation of the $\bar{y}_i$ may be different for different trial periods, and a measurement appropriate for one form would hardly apply to another.

Now the situation actually met in the practical investigation of economic series is generally one which leads to the more complicated wave-form. The typical average cycle in an economic record does not ordinarily resemble closely a simple sine curve. Usually there is a distortion which might be described as a skewing of the sine curve: the time interval from low to high considerably exceeds that from high to low. Moreover, the typical cycle generally exhibits irregularities which may be regarded as supplementary fluctuations—in some instances quite marked—superimposed upon the fundamental wave. Both of these characteristics of the typical economic cycle render a description of the cycle by the first periodic term in the Fourier series insufficient and require that some or all of the remaining terms of the expansion be found.

The calculation of the coefficients in the complete Fourier expansion is, however, somewhat tedious; and, as suggested above, the coefficients when found do not yield a simple or a sufficient criterion of selection among the several trial periods. Furthermore, the description of the cycle by a complete Fourier expansion is unlikely to be interesting

for its own sake, partly because the significance of the individual terms in the expansion is difficult for the mind to grasp, and partly because we are still far from accepting on *a priori* grounds the principle that economic fluctuation is compounded of a group of sine curves.

II

We adopt now a somewhat different point of view. Let it be required to find the correlation between the given data and a sine function. To fix our ideas, suppose a particular item of the original data is represented by Y_{ij} and falls in the ith column and the jth row of the rectangular tabulation scheme. Suppose further that i ranges from 1 to p, and that j ranges from 1 to s, and that $N=sp$. Finally, take as the sine function:

$$F=\sin\left(\frac{2\pi}{p}i+\alpha\right) \tag{3}$$

where, in the notation of Section I,

$$\alpha=\sin^{-1}\frac{A}{R}=\cos^{-1}\frac{B}{R}. \tag{4}$$

The desired correlation is then:

$$r=\frac{\frac{1}{N}\Sigma y_{ij}\sin\left(\frac{2\pi}{p}i+\alpha\right)}{\sigma_y\sigma_F} \tag{5}$$

where y_{ij} is the deviation of Y_{ij} from the mean of the N original items.

We proceed now to evaluate the several factors in (5).

$$\begin{aligned}\sigma_y^2&=\frac{1}{N}\Sigma y_{ij}^2\\&=\frac{1}{N}\Sigma\Sigma\left\{(y_{ij}-\bar{y}_i)^2+2\bar{y}_i y_{ij}-\bar{y}_i^2\right\}\\&=\frac{S}{N}\Sigma\left\{\sigma_i^2+2\bar{y}_i^2-\bar{y}_i^2\right\}\\&=\bar{\sigma}_i+\sigma_{\bar{y}}^2\end{aligned} \tag{6}$$

where σ_i is the standard deviation of the items in the ith column, and $\bar{\sigma}_i$ is the mean value of the σ_i, and $\sigma_{\bar{y}}$ is the standard deviation of the column averages.

$$\sigma_F^2=\frac{1}{p}\Sigma\sin^2\left(\frac{2\pi}{p}i+\alpha\right)$$

$$= \frac{1}{2p} \Sigma_i \left(1 - \cos\left[\frac{4\pi}{p} i + 2a\right]\right)$$

$$= \frac{1}{2} \tag{7}$$

$$\frac{1}{N} \Sigma_{ij} y_{ij} \sin\left(\frac{2\pi}{p} i + a\right) = \frac{1}{p} \Sigma_i \bar{y}_i \sin\left(\frac{2\pi}{p} i + a\right) = \frac{1}{2} R \tag{8}$$

by use of (1), (2), (4) and the relation:

$$R = A \sin a + B \cos a.$$

Hence, from (5), (7), (8):

$$r = \frac{\frac{1}{2}R}{\frac{1}{\sqrt{2}}\sigma_y}$$

$$= \frac{R}{\sqrt{2}\sigma_y}. \tag{9}$$

We observe, as might have been foreseen, that the R of the periodogram analysis is closely connected with the correlation coefficient. Except for the divisor $\sqrt{2}$, the coefficient equals the ratio of R by σ_y. Strictly, if we were using the value of r as the criterion of selection among a range of trial periods, we should divide the value of R calculated for each trial period by the corresponding value of σ_y. An obvious exception would occur in comparing two trial periods for which identically the same N items were used in the analysis; and, in general, the slight actual variations in the N items would scarcely be reflected by large variations in σ_y. Hence it is doubtful if we should be led by (9) to make a change in the practical procedure of the periodogram analysis.

From a theoretical point of view, however, the above study of r is suggestive of important considerations bearing upon the use of R, and, indeed, of the periodogram method. The fact that R is essentially a correlation coefficient means that for it to be significant as a comparative measure, regression must be closely linear. Obviously, the cases in which the first periodic term of the Fourier expansion does not adequately describe the $\bar{y}_i$ are precisely the cases in which the regression of the $\bar{y}_i$ relative to F is not linear.

We are therefore led again to the conclusion that the periodogram method does not furnish a satisfactory means of analyzing the typical economic cycle in which the fluctuations depart notably from a simple sine curve.

III

The obvious alternative procedure, if we continue to regard the problem as one in correlation, is to examine the correlation ratio rather than the correlation coefficient. Although the treatment would be in closer conformity with that of Section II if we determined the correlation ratio of y_{ij} on F, the analysis is somewhat more direct and the results more significant for the practical problem in hand—the selection of a period from a range of trial periods—if we study the correlation ratio of y_{ij} with i.

We define the correlation ratio of y_{ij} on i by η, where

$$\eta = \sigma_{\bar{y}}/\sigma_y. \tag{10}$$

This ratio η measures the average tendency of the items of the several columns to conform to their respective means: if η is large, it appears from (6) that $\bar{\sigma}_i/\sigma_y$ is correspondingly small and the items y_{ij} do not deviate greatly on the average from $\bar{y}_i$. It would seem that this is precisely the situation to be desired in the case of the particular trial period selected as the best approximation to the true period. What is sought is not a trial period p for which the amplitude R of the simple sine curve of best fit (which may be poor indeed) is greatest, but rather a trial period p for which the column averages $\bar{y}_i$ are most nearly typical of the items y_{ij} in every row j. A good criterion for this condition of high conformity to type is furnished by η.

The calculation of η from (9) is exceedingly simple, much simpler in fact than that of R. One carries the regular process of the periodogram analysis through to the stage of finding the $\bar{y}_i$ for each trial period p. Then, for each p, the value of σ_y for the N items y_{ij} actually used in studying that period and the value of $\sigma_{\bar{y}}$ are computed; and the ratio of the second result by the first is η for the p in question. The p which yields the largest η is chosen as the tentative value of the unknown period. It should be remarked that all the numerical work will be simplified if the y_{ij} are entered as deviations from a value not far from their general mean, and the alterations involved in the "arbitrary origin" device should be made in the several formulas. Other aids in the numerical work will probably suggest themselves, but the labor of computation is at worst not unduly heavy.

IV

(a) We observed at the beginning of Section III that the correlation ratio η selected for the periodicity problem pertains to y_{ij} and i; whereas the correlation coefficient r, which is involved in the amplitude

R^2 of the periodogram, measures the correspondence between y_{ij} and the F function defined by (3). It is therefore not possible to take

$$\eta^2 - r^2$$

as the measure of linearity of regression and thus find the goodness of fit of the first periodic term in the Fourier expansion to our data.

For such a purpose it would be necessary to find η as the correlation ratio of y_{ij} on F, which involves classifying the original data according to values of F—a sine function of the time—rather than of i—the time itself. Reference to (3) shows:

$$i = \frac{p}{2\pi}(\sin^{-1} F - \alpha) \tag{11}$$

and (11) should enable us to determine intervals of i which correspond to uniform intervals of F, and hence to regroup the given data according to uniform class-intervals in F instead of in i. For each value F_k belonging to the center of one of these new class-intervals we should have then an array of y_{kj} including items y_{ij} from one or more of the original i—arrays. With the regrouping accomplished, the η for y on F could be calculated.

There is, however, a serious practical obstacle to this regrouping. Our data come classified relative to time; and, although they may pertain to an economic movement which is essentially continuous, they are themselves discontinuous. The breaking of one of the original i—arrays and its distribution among two or more of the new F—arrays, although theoretically possible if certain assumptions are made as to the form of the actual continuous fluctuation, would be difficult in practice.

Fortunately, it is not essential to a determination of the goodness of fit of the Fourier term that we find η for y on F. As soon as R and α are found, in the usual periodogram analysis, the ordinates of the fitted sine curve for all values of i can be calculated. Let these ordinates be represented by z_i, and then

$$\frac{1}{R^2}\Sigma_i(z_i - \bar{y}_i)^2$$

may be used to indicate the goodness of fit.

(b) In regard to the goodness of fit of the wave-form selected by the method of Section III, it may be noted at once that η itself serves to indicate how typical this cycle is. In fact, by (6) and (10),

$$\bar{\sigma}_i^2 = \sigma_y^2(1 - \eta^2) \tag{12}$$

and this shows that, for a given σ_y, the nearer η is to unity the smaller will $\bar{\sigma}_i$ be, and on the average the closer will be the fit of the $\bar{y}_i$ to the y_{ij}.

If one wishes to examine in detail the closeness of fit of each $\bar{y}_i$ to the corresponding y_{ij}, one should calculate the several values of σ_i. This involves getting the sums of the squares of the items in each column of the rectangular tabulation scheme, but does not add greatly to the arithmetic of the problem. As soon as σ_i is known for all i one can set up the function:

$$\lambda_i = \frac{\sigma_i}{\sigma_y} \tag{13}$$

the graph of which is the scedastic curve of the series.

If the total time interval covered by the problem is not long, the number, s, of items in a particular i— array will be small and σ_i will be subject to marked fluctuations. On the other hand, even if s is not small, it is possible that the nature of the cycle being investigated is such that certain portions of it are more precisely typical than others; and, for such a case, (13) will yield λ_i which vary over a considerable range.

When the σ_i are calculated in this manner, it is possible to get $\bar{\sigma}_i$ directly from them. With $\sigma_{\bar{y}}$ known, σ_y can then be obtained from (6). In general, however, one will not trouble to get the σ_i except for those values of p suggested by the preliminary analysis as tentative periods.

(c) The part played by the standard deviation in the periodicity analysis suggests a criticism of the validity of that measure of dispersions for a time series. It has often been pointed out that the standard deviation is not so good an index of dispersion in an ordered series as in a simple frequency series. The fact of ordering renders it probable that adjacent items of an ordered series will be subject to similar fluctuations. This condition must be borne in mind in the use of the above formulas: wherever a σ enters it pertains to a series ordered in time. This is obviously true for σ_y, and a brief consideration shows that it holds indirectly for σ_i, $\bar{\sigma}_i$, and $\sigma_{\bar{y}}$ also. Search for a substitute for σ, which shall be free from its objectionable features so far as ordered series are concerned and yet retain its mathematical advantages, has not been very successful. Although σ is not ideal for the purpose, the actual results which it yields in practice appear sufficiently reliable to justify its use in the periodicity problem until it shall have been replaced by a dispersion index clearly better for time series.

(d) Mention has been made of the apparent skewing of the typical

economic cycle—the tendency for the decline to be more abrupt than the rise. This may account in considerable measure for the failure of a simple sine curve to furnish a good fit to the $\bar{y}_i$. We saw in Section I that it was not desirable to replace the simple sine term by a complete Fourier expansion.

Suppose, on the other hand, we seek a periodic function of a single term which shall have the desired property—the time interval from minimum to maximum greater than that from maximum to minimum. Such a function is determined by:

$$y = \mu \sin a(x - y \tan \theta) \tag{14}$$

where μ, a, and θ are constants, and x depends linearly upon the time.

The maximum and minimum points of (14) in the first wave-length are:

$$(\mu \tan \theta + \frac{\pi}{2a}, \mu), (-\mu \tan \theta + \frac{\pi}{2a}, -\mu)$$

and it is clear that the $x-$ intervals from minimum to maximum and from maximum to minimum are respectively:

$$x_1 = \frac{\pi}{a} + 2\mu \tan \theta, \; x_2 = \frac{\pi}{a} - 2\mu \tan \theta \tag{15}$$

and differ from each other by $4\mu \tan \theta$.

In the practical problem this difference $4\mu \tan \theta$ will be small compared to p; and, as ap is $2\pi \cos \theta$, we may replace cos (aytan θ) and sin (aytan θ) by 1 and aytan θ respectively in an approximation to the first order. Then (14) is:

$$\begin{aligned} y &= \mu \frac{\sin ax}{1 + a\mu \tan \theta \cos ax} \\ &= \mu \sin ax \{1 - a\mu \tan \theta \cos ax\} \\ &= \mu \sin ax - \frac{\mu^2 a}{2} \tan \theta \sin 2ax \end{aligned} \tag{16}$$

and this is the expression by which we should replace the F of (3) in order to handle the correlation problem for the skewed curve. It is understood, of course, that x is $i + \beta$, where β corresponds to the α of (3), and a is as given above.

Now (16) is analogous in form to the first two terms of the Fourier expansion; for, although the coefficient of the second term in (16) bears a definite relation to that of the first, that relation depends upon a tan θ. In practice, if the coefficients were determined by the Fourier

method we could get tan θ from these coefficients rather than *vice versa*. We should note of course that the phase angles of the two terms in (16) are not independent as they would be in the Fourier development. In general, however, we may say that the fitting of our approximate form of the skewed wave would hardly yield different results than those obtainable by getting the second Fourier term. A different analysis by use of equation (14) may be possible, but there is no obvious simple means of attacking it.

(e) The relation between the R of the periodogram and the correlation coefficient suggests the consideration of the problem of lag from a different point of view. If we are seeking the correlation of the cycles of one economic record with those of another, it is altogether likely that the highest correlation will be found when pairs of items differing in time by a fixed interval called the lag are associated with each other. The determination of the lag is often difficult because the actual lag between successive pairs of cycles is not strictly constant and therefore not readily assessed by inspection.

If, now, we assume that each series has been analyzed by the above method, a comparison of the value of a for one series with that for the other will show the lag. The procedure would be to determine, by the η method, p for each series; and then to find, by the Fourier method, a for each series. If the origin of time is identical for the two series, the lag is the difference between the values of a. This method would scarcely be applicable unless the typical cycle for each series is moderately close to a sine curve, and in any case the labor is considerable unless the periodicity analysis is needed for another purpose.

V

We present now a periodicity analysis of a particular economic series by each method. The series chosen is the record[1] of the rate of interest on 60–90 day commercial paper from July, 1899, to June, 1914. Six trial periods were examined—p ranging from 40 to 45—and the total interval used in the rectangular tabulation scheme covered the months May, 1900, to August, 1913, for $p=40$, and was increased symmetrically for the other trial periods. For each p, s was constantly 4. The column averages for the six trial periods are shown in Table I.

The calculations were carried out by use of formulas given in Sections I and III; and the resulting values for A, B, R^2, a, σ_y^2, $\sigma_{\bar{y}}^2$ and η are given in Table II. The results for both R^2 and η suggest that p might reasonably be taken as 42 or 43, and R^2 alone would probably not exclude 41 from the choice. If a chart be made of the column averages

[1] *Rev. Econom. Stat.*, Jan., 1923, p. 28.

with the fitted sine curve, for the 42-months and 43-months trial periods, it will be seen that the fit is fairly good in both cases. The calculation of the mean square deviation might enable us to say that the fit is slightly better in one case than the other, but it is doubtful if it would give a more reliable conclusion than that afforded by an inspection of the charts.

We have here a series in which the cycle is fairly close to a sine curve of period 42–43 months. This does not justify us in confidently predicting that the interest rate will continue to move in this manner. It warrants the conclusion that for the interval studied the average cycle had approximately the form we have found. Moreover, our calculations indicate that for this particular series the η yields quite as good a criterion for the period as does R^2; and, for series which conform less well to the sine curve, we may expect η to be the more reliable.

TABLE I

COLUMN AVERAGES FOR VALUES OF p FROM 40 TO 45

Unit: 1/00 per centum

40	41	42	43	44	45
10	45	48	53	34	14
7	39	43	71	45	14
40	24	38	65	65	22
35	28	39	46	60	14
6	23	10	21	13	23
21	1	—9	—9	—6	30
9	—12	—33	—20	12	4
—7	—1	—17	—32	00	7
—31	—16	—7	—18	—29	31
—40	—10	—4	—11	—20	00
—34	—31	—13	—10	00	—1
—12	—29	00	19	11	23
—45	—60	—60	—39	—23	—32
—73	—72	—80	—53	—50	—73
—54	—45	—45	—53	—63	—46
—66	—76	—72	—93	—59	—79
—66	—78	—102	—82	—96	—101
—88	—100	—88	—109	—113	—78
—92	—71	—103	—106	—105	—92
—72	—98	—108	—108	—100	—110
—94	—91	—76	—95	—99	—101
—111	—71	—79	—83	—104	—87
—53	—91	—77	—77	—80	—79
—83	—96	—86	—75	—67	—68
—71	—94	—80	—70	—45	—52
—77	—69	—82	—46	—54	—57
—42	—51	—47	—49	—48	—45
—21	—21	—29	—52	—34	—77
—8	—5	—16	—32	—69	—68
4	—29	—13	—49	—72	—33
3	2	—57	—34	—22	—42
37	—3	—30	—19	—4	—25
—17	8	23	—15	—22	6
1	53	40	12	10	16
46	25	38	19	26	16
45	19	56	61	34	15
35	54	35	44	38	9
43	56	32	43	33	4
72	39	59	33	32	27
74	47	48	43	61	18
	54	33	57	24	9
		46	20	—5	30
			6	23	14
				16	10
					30

TABLE II

VALUES OF THE CONSTANTS FOR THE VARIOUS TRIAL PERIODS

p	40	41	42	43	44	45
A	.64	.66	.66	.66	.62	.55
B	—.16	—.18	—.21	—.20	—.22	—.13
R^2	.43	.47	.48	.47	.43	.32
α	—76	—75	—72	—73	—70	—76
σ_y^2	.84	.83	.82	.81	.81	.81
$\sigma_{\bar{y}}^2$	.24	.26	.27	.27	.25	.20
η	.54	.56	.58	.58	.55	.49

Units: For A and B, one per centum; for R^2, σ_y^2, and $\sigma_{\bar{y}}^2$, the square of one per centum; for α, one degree.

NOTES

MONGOLIAN AMENTIA AND THE AGES OF THE MOTHERS AT THE BIRTHS OF THE AMENTS

By Hornell Hart, *University of Iowa*

While searching recently for illustrative material for a course on statistical methods in social research, I came across some data collected by Goddard showing the age distributions of the parents of feebleminded children of the Mongolian type in certain institutions, at the dates when the children were born.[1] After citing the data, Goddard says: "The number of children born when the mother is 40 is much higher than at any other age. This may or may not be significant."

It is so easy to show that the age distribution of the mothers at the time these children were born is significant, and the results of the analysis have so definite a bearing on the etiology of this type of amentia and on that of other abnormalities, that the material furnishes an excellent illustration of the importance of statistical technique in social sciences.

The abnormality of the age distribution of the mothers of these Mongolian feebleminded individuals at the times of their births may be shown by a comparison with the age distribution of the mothers of ordinary babies at the times of their births. The most readily available data of this sort are found in the infant mortality studies of the United States Children's Bureau, where the age distributions of the mothers of all babies born in certain cities in selected years are tabulated.[2] The data from five of these cities (Akron, Brockton, Johnstown, Manchester, and Waterbury) are used in Table I.

This table shows that the tendency of such Mongolian aments as those cited by Goddard to appear in such a population as that found in five typical American cities is 8.02 times the average rate when the mother is 40 years of age or over at the child's birth, and is only 0.34 times the average rate when the mother is 20 to 24 years old at the baby's birth. In other words, the Mongolian amentia rate is over 23 times as great among persons born of mothers 40 years of age or over as it is among persons born of mothers 20 to 24. The probability that the observed deviations from the average rate are due to chance, in

[1] Goddard, H. H., *Feeblemindedness: Its Causes and Consequences*. New York. 1916. Pp. 451-2.
[2] United States Department of Labor, The Children's Bureau. Infant Mortality Series.

TABLE I

PERCENTAGE DISTRIBUTION, ACCORDING TO THE AGE OF THE MOTHER AT THE CHILD'S BIRTH, OF MONGOLIAN AMENTS CITED BY GODDARD, IN COMPARISON WITH THE BABIES BORN IN SPECIFIED YEARS IN FIVE AMERICAN CITIES, AS REPORTED BY THE UNITED STATES CHILDREN'S BUREAU

Age of mother	Percentage distributions of		Relative ratio of Mongolian amentia
	Mongolian aments	Representative babies	
15-19	3.4	5.7	0.60
20-24	10.2	29.8	0.34
25-29	13.6	28.9	0.47
30-34	11.9	19.1	0.62
35-39	24.0	11.9	2.02
40-44	30.5	....	....
45-49	5.4	....	....
50-54	0.7	....	....
55-60	0.3	....	....
40-59	36.9	4.6	8.02
All ages	100.0	100.0	1.00
15-34	39.1	83.5	0.47
35-59	60.9	16.5	3.69
Average age of mother	38.6 yrs.	30.7 yrs.	
Number of cases	295	8,995	

the sense that they would disappear if enough cases were collected by the methods used in collecting the data involved in the above comparison, is infinitesimal, as determined by Yule's formula.[1]

It will be noted that the relative ratio of Mongolian amentia in children whose mothers were under 20 at the time of their birth is higher than that among children whose mothers were 20 to 24. If only mothers under 25 are considered, and the percentage age distributions of the mothers are compared, it appears that the tendency of the mothers of Mongolians in this age group to be younger than the mothers of average children is greater than is likely to be exceeded by chance once in 50 such samples.

Another striking comparison is afforded by Cattell's data on the ages of the parents of American men of science at their births.[2] He finds that the average age of the mothers of 865 scientists at their births was $29.70 \pm .15$ years—a year younger than the average of the mothers of babies studied by the Children's Bureau, and nine years younger than that of the mothers of the Mongolian aments. For the 64 most distinguished scientists the mothers averaged $28.37 \pm .50$ years at their birth, or about a year and a half younger than the mothers of

[1] Yule, G. Udney, *An Introduction to the Theory of Statistics*. London. 1919. P. 269.
[2] *American Men of Science*, 1921, p. 801.

less distinguished scientists, and about two and one-third years younger than those of the Children's Bureau babies.

The data relative to the ages of the mothers of men of science thus emphasize the abnormality of the age distribution of the mothers of the Mongolian idiots, and suggest that men of superior intelligence tend to be born more nearly at their mothers' physiological prime than is the average man.

In comparison with Goddard's discussion of the subject it is interesting to note the statements of two other writers on mental defectiveness. Barr writes: [1]

> Again old age in parents has been considered a cause of Mongolianism, but I know of but one instance in which this could be ascribed, the father being 75, and the mother 48 years of age, the average age of parents being 36 for the fathers and 33 for the mothers. It is also said that last-born children are apt to be Mongolians, but I find of this but one case. Indeed, I do not find data sufficiently reliable to affirm any one cause above another in the production of this type.

It will be noted that the average age of the mothers cited by Barr is five or six years lower than that shown by Goddard's data, and is two or three years older than that shown for the normal babies studied by the Children's Bureau. In the absence of information as to the number of cases studied by Barr and the standard deviation of the ages of the mothers in his group, it is impossible to tell how reliable his figure is. Data courteously furnished the writer by Dr. George Mogridge, Superintendent of the Institution for Feebleminded at Glenwood, Iowa, relative to 43 Mongolians, show an age distribution for the mothers quite similar to that in Goddard's data.

Tredgold says: [2]

> In nearly all my cases there was a neuropathic family history, and frequently a strong tubercular taint; but over and beyond this what I did frequently find (and what I think may possibly be the factor common to this type) was a history pointing to a condition of uterine exhaustion or ill-health of the mother during gestation. Many of the patients were the later born of a large family, often numbering as many as ten or twelve, and where this was not the case there was usually a state of severe physical prostration of the mother during the gestation period.

A striking comparison with the data relative to Mongolians is furnished by the findings of Antonio Marro [3] (Table II).

Among criminals the observed excessive proportion born of immature mothers would be exceeded by chance once in about 1,000 such samples, while the observed excessive proportion born of mothers over 36

[1] Barr, Martin W., *Mental Defectives*. Philadelphia. 1904. P. 83.

[2] Tredgold, A. F., *Mental Deficiency*. New York. 1914. P. 213.

[3] Marro, Antonio, "The Influence of the Age of the Parents upon the Psycho-physical Characteristics of the Children." Problems in Eugenics. The Eugenics Education Society. London. 1912. Pp. 118–36.

TABLE II

PERCENTAGE DISTRIBUTIONS, ACCORDING TO AGE OF MOTHER AT CHILD'S BIRTH, OF PERSONS EXAMINED IN MILAN, ITALY, BY ANTONIO MARRO

Age of mother	Normal persons	Criminals	Relative ratio of criminality	Insane persons	Relative ratio of insanity
-20	12.8	18.2	1.42	20.0	1.56
21-36	76.4	63.7	0.83	58.8	0.77
37-	10.7	17.9	1.67	21.1	1.97
All ages	99.9	99.8	1.00	99.9	1.00
Number of cases	1,765	456		85	

years of age would be exceeded by chance about once in 300,000 such samples. Among the insane persons the observed excess of persons born of mothers under 21 would be exceeded about once in 15 such samples (and hence might be due to chance), while the excessive proportion of those born of mothers over 36 years old would be exceeded by chance about once in 200 such samples. The apparent conclusion is that criminality tends to be about twice as frequent and insanity about two and a half times as frequent among persons born of old mothers as among those born of mothers in the mid-period of fertility. Marro's data suggest that while criminality is associated also with the age of the father, the age of the mother is the determining factor.

Woodbury shows that the infant mortality rate for children born of mothers under 20 or over 40 is more than one-third higher than for children born of mothers 25 to 29 years of age.[1]

The data presented above suggest three different possible hypotheses. First is the possibility that the individuals with defective germ plasm or with especially unfavorable environments tend to marry earlier and to bear children longer than other individuals. This explanation is fairly plausible in connection with the infant mortality data, but it does not account for the absolute excess of Mongolians with mothers who were 40 to 44 years of age at the child's birth. A second hypothesis is that repeated pregnancies have exhausted the mother (as Tredgold suggests), and hence the prenatal conditions are unfavorable. Such a theory would not necessarily involve any supposition that the germ plasm of later pregnancies was inferior to that of earlier ones. The third hypothesis is that the quality of the germ plasm produced varies with the age of the mother, reaching its optimum in early maturity. The use of partial correlation or partial association methods, in conjunction with technical researches in genetics, can determine which of these hypotheses is correct.

[1] This JOURNAL, September, 1922, p. 375.

LONG TERM PRICE TRENDS

By Lawrence K. Frank

Of late there has been considerable speculation upon the probable future trend of prices. Although the writer can contribute nothing by way of prophecy, the tentative hypothesis regarding long term price trends offered in this paper may prove of interest to those concerned with that question.

Two points of interpretation or assumption are involved in the following discussion. They will be stated first as preliminary to the statement of the hypothesis:

I. The long term or secular trend in economic activities is generated by cyclical fluctuations. To illustrate, the production of pig iron in the United States fluctuates in magnitude from year to year, but over a series of years there may be discerned a long term growth. This long term growth occurs by reason of the fact that the production of pig iron, in successive periods of expansion, reaches a progressively larger volume and, in general, is not reduced, in successive periods of contraction or depression, to as low a level as in the previous period of depression. It is obvious that if the production of pig iron is curtailed in successive periods of depression to 50 per cent of plant capacity and if the plant capacity is being increased in each period of expansion, the reduction in volume of production will be progressively less. In so far as the volume of production is increased in periods of expansion above the previous high figure, the secular trend will be upward. Consequently we may say that the cyclical fluctuations in magnitude of pig iron production give rise to the long term growth in pig iron production. More generally we may say that the long term or secular trend is generated or produced by cyclical fluctuations.[1]

II. Cyclical fluctuations in prices are indexes of variations in the volume of purchases. To illustrate, it has been shown[2] that the cyclical fluctuations in the prices of commodities usually occur in the following order and magnitude: In a period of expansion, raw materials (non-agricultural) rise first and most, then semi-finished goods, then finished goods at wholesale, and last and least, retail goods; moreover, in a period of contraction, the fall of prices is in the same order and

[1] It will be seen that if growth is not a continuous process it must be cyclical, *i. e.*, an affair of increase and recession or quiescence in which the successive increases are progressively larger in order to compensate for the recessions and pauses and to add the increments necessary for growth. Growth by cycles is the ordinary and usual case in biology. See T. W. Robertson, *Principles of Biochemistry*, page 471 *et seq.* for examples of cyclical growth in organisms and in chemical reactions.

[2] W. C. Mitchell, *Business Cycles*, page 93 *et seq.*

magnitude as the prior rise. Furthermore, the rise in the price of producer's goods is greater than the rise in the price of consumer's goods; and the rise of all prices is more irregular than the fall of all prices.[1]

From the available data on production,[2] it appears that in a period of expansion the increase in magnitude of production is greatest for raw materials (non-agricultural), somewhat less for semi-finished, and least for finished goods, while in a period of contraction the decrease in magnitude of production is in the same order. Moreover, the volume of retail trade fluctuates less than the production of finished goods, both in periods of expansion and in periods of contraction.

The similarity in order and magnitude of cyclical change in prices and production suggests that they are both concomitants of changes in the volume of purchases; for it seems clear that the increases and decreases in production (non-agricultural) are, however roughly, responses to the orders or purchases of buyers. Even where goods are not made to order but are sold from stocks on hand, the volume of production will vary with the changes made in those stocks by sales. And it appears equally clear that the rise and the fall in prices are directly related to the changes in volume of purchasing, since increasing purchases and increasing prices and decreasing purchases and decreasing prices occur together. Moreover, the differential price changes, which we have already discussed, appear fairly clearly to be responses to changes in volume of purchasing.

The foregoing is purely inference and is subject to correction and refinement. It is explicitly given here as merely an assumption or interpretation underlying the hypothesis of long term price trends.[3]

When we examine the long term trend in prices in the light of the foregoing assumptions the situation appears to be this: In a long term period of falling prices, such as that which terminated in 1897 (approximately), there are found the usual cyclical fluctuations, but in general the price rise in successive periods of expansion is progressively less, while in successive periods of contraction the price decline is progressively greater. Per contra, in the period of rising prices since 1897, the price rise in successive periods of expansion was progressively greater, while the price decline in successive periods of contraction was progressively less. Certain exceptions to these statements may be noted,

[1] *Ibid.*, page 109 *et seq.*

[2] See production indexes by E. E. Day, W. I. King, Carl Snyder, Walter Stewart, and material appearing in the *Review of Economic Statistics*, the Federal Reserve Board Bulletin, and the *Review of Credit and Business Conditions* of the Federal Reserve Bank of New York for the past two or three years.

[3] For further discussion of cyclical price changes and production changes and their significance, see the writer's paper on "Business Cycles," *Quar. Jour. of Economics*, August, 1923.

but they do not detract from the general proposition about long term trends.

From the earlier discussion of changes in prices as indexes of changes in volume of purchases the price changes described above may be interpreted as follows: In a long term period of falling prices, the short upward swings in prices indicate that producers and merchants do not increase the volume of their purchases as much in each period of expansion as in the previous period of expansion; in actual business practice this means that they do not build up such large stocks of the goods they use or sell. Likewise, in the short downward swings in prices, producers and merchants draw down their stocks to lower amounts, thus further decreasing the volume of their purchases or refraining from purchases for a longer period. It has frequently been noted that when prices are falling over a long period, the periods of expansion (and of price rise) tend to be short and the periods of contraction (and of price decline) tend to be long. When prices are rising over a long period the reverse is true.

In a long period of rising prices, producers and merchants in successive periods of expansion increase their purchases progressively more; that is, they build up larger stocks of the goods they use or sell. In successive periods of contraction, producers and merchants do not permit their stocks to be drawn down so much as in preceding periods of contraction. Or to put it another way, they begin to increase their purchases earlier in successive periods of contraction, for it is a well known fact that the fall in prices in depression is halted when producers and merchants begin to buy. Likewise, in expansion the rise of prices continues as long as buying or purchasing grows in volume.

It should be noted that these up and down swings in purchases and stock accumulations may occur even though there is a fairly continuous steady growth in population and final consumption (which is not quite so steady). For, while population growth is a physical fact, purchases and stock accumulations are pecuniary operations which recurrently increase and decrease in volume and thereby produce a resultant secular growth roughly parallel to population growth. Also it should be noted that these variations in volume of purchases and stock accumulations may occur even though production is growing, as is pointed out earlier in the discussion of pig iron production. Consequently the change from one long term price trend to another may occur without corresponding changes in population, consumption or production trends.

In general, therefore, the magnitude of cyclical fluctuations in prices appears to be correlated with the extent to which producers and merchants build up their stocks of goods in periods of expansion or sell and

reduce those stocks in periods of contraction. And the long term trend of prices appears to be a resultant of the variations in the extent of accumulation and depletion of stocks of goods.

Since the extent of accumulation of stocks of goods (which is a form of speculation) in periods of expansion and the extent of stock depletions in periods of contraction are intimately related to the changes in credit conditions, it may well be that the connection between changes in the production of gold and the long term trend of prices is to be found in the operations herein discussed. Obviously, if an increasing supply of gold or any other circumstance provides the basis for a larger supply of commercial credits (short term bank loans) and the expansion of these credits occurs by reason of the fact that producers and merchants are building up stocks of goods, then the connection sought appears manifest.

The theory or hypothesis herein proposed is not offered as a final solution of the problem; it merely suggests the economic activity or behavior which relates changes in credits and in gold supplies to the long term trend in prices. Incidentally it calls attention to the fact that any circumstance or event which affects the volume of purchasing by producers and merchants (*e. g.*, political agitation or proposed monetary changes) may influence the cyclical fluctuations in prices and, hence, the long term trends which are generated by the cyclical fluctuations.[1]

With respect to the prices of agricultural commodities, the theory of their cyclical changes would be as follows: When the volume of purchases of those commodities is increasing (in a period of expansion) the prices will rise, but the extent of the rise will be conditioned by the size of the crop yields: a large crop yield will retard the rise and a small crop yield will accelerate the rise. When the volume of purchases is decreasing (in a period of contraction) the prices will fall, but, again, a large crop yield will accelerate the price fall and a small crop yield will retard the price fall.[2] If the production of agricultural commodities is increasing faster or more slowly than the growth of purchases or consumption, the long term trend in prices of agricultural commodities will be accordingly down or up.

In conclusion it may be added that most authorities attribute the long term price trend to the trend in production of gold, but they do not explicitly state how the gold production is translated into price changes. It is the purpose of the foregoing discussion to point out one

[1] It is an interesting question whether the cyclical fluctuations which generate the secular trends are not themselves generated by seasonal fluctuations.

[2] Exports may alter the situation very quickly by suddenly increasing the volume of purchases even though business conditions are still unfavorable. *Cf.* 1896.

way in which that translation may occur and to suggest further study of the activity or the behavior of men as important in accounting for pecuniary changes.

SPURIOUS CORRELATION APPLIED TO URN SCHEMATA

By John Rogers Musselman

The term "urn schemata" is used to mean any game of pure chance, such as drawing balls from an urn, or tossing coins or throwing dice. By its use many theorems in the theory of probability are more simply and precisely stated. Its importance in mathematical statistics is well known. It is used in the theory of sampling and should serve as a basis for the development of the theory of the correlation coefficient.

Weldon [1] and Darbishire [2] have performed experiments with dice and have shown empirically facts about the correlation coefficient. One of the problems of Rietz,[3] solved from the standpoint of theoretical probabilities, if clothed in different terminology will prove the empirical results of Darbishire. It is the purpose of this article to derive these results from a different point of view, with much more ease and simplicity—one which permits of extension to the subject of partial correlation coefficients.

Theorem I. Throw s dice, of which t are colored red and the others white, and count the number of sixes appearing face up. Let the red dice lie, throw the white again, and count the number of sixes on both the red and white dice. If m pairs of throws be made, and a correlation table be set up between the number of sixes appearing on the first throw and on the second throw, the correlation coefficient is t/s.

In games of pure chance such as this there can be no inherent correlation, and some scheme, therefore, must always be introduced to obtain correlation between the two sets of throws; *e. g.*, in the above theorem the red dice were used in both throws. Hence the correlation arising is entirely spurious correlation. If it is possible to measure this spurious correlation independently, we shall know the correlation coefficient. This can be done by the following method.

Let us consider a set of variables z made of two subsets x and y which can take n values.

$$z = x + y.$$

[1] Weldon, *Lectures of the Method of Science*. Edited by T. B. Strong, Oxford, 1906.

[2] Darbishire, *Memoirs and Proceedings of the Manchester Literary and Philosophical Society*, Vol. 51, No. 16, 1907.

[3] Rietz, *Annals of Mathematics*, Second Series, Vol. 21, No. 4, 1920. Five problems are discussed by Rietz, two of which come under this theorem. In his other problems the regression curves are not both linear and therefore the correlation coefficient cannot be used.

The mean and standard deviation of the set of variables z can be expressed in terms of the means and standard deviations of the variables x and y.[1]

$$z_m = x_m + y_m,$$
$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2r_{xy}\sigma_x\sigma_y.$$

If we have two sets of variables, $z = x+y$, $z' = x'+y'$, each assuming n values, the correlation coefficient $r_{zz'}$ can also be expressed in terms of the coefficients of the subsets, for

$$(1)\quad r_{zz'} = \frac{\Sigma zz'}{N\sigma_z\sigma_{z'}} = \frac{\Sigma(x+y)(x'+y')}{N\sigma_z\sigma_{z'}}$$
$$= \frac{r_{xx'}\sigma_x\sigma_{x'} + r_{xy'}\sigma_x\sigma_{y'} + r_{x'y}\sigma_{x'}\sigma_y + r_{yy'}\sigma_y\sigma_{y'}}{\sqrt{\sigma^2{}_x\sigma^2{}_y + 2r_{xy}\sigma_x\sigma_y}\cdot\sqrt{\sigma^2{}_{x'}\sigma^2{}_{y'} + 2r_{x'y'}\sigma_{x'}\sigma_{y'}}}.$$

If no correlation exists between the subsets x, y, x', y', that is, if $r_{xx'} = r_{xy'} = r_{x'y} = r_{yy'} = 0$, then $r_{zz'} = 0$ also.
However, if we put $x' = x$ in (1) then

$$(2)\qquad r_{zz'} = \frac{\sigma_x^2 + r_{xy'}\sigma_x\sigma_{y'} + r_{xy}\sigma_x\sigma_y + r_{yy'}\sigma_y\sigma_{y'}}{\sqrt{\sigma_x^2 + \sigma_y^2 + 2r_{xy}\sigma_x\sigma_y}\sqrt{\sigma_x^2\sigma_{y'}^2 + 2r_{xy'}\sigma_x\sigma_{y'}}}.$$

And if we suppose as before that no correlation exists between x, y, y'; that is, if $r_{xy'} = r_{xy} = r_{yy'} = 0$, then $r_{zz'} \neq 0$, but takes the value

$$(3)\qquad \rho = \frac{\sigma_x^2}{\sqrt{\sigma_x^2 + \sigma_y^2}\sqrt{\sigma_x^2 + \sigma_{y'}^2}}.$$

Hence, despite the fact that there is no correlation between the subsets of z and z', nevertheless from the mathematical definition of correlation $r_{zz'} \neq 0$. This is termed spurious correlation and is due simply to the variables z and z' containing the common subset x.

The application of this to the proof of the theorem is evident. The correlation there is entirely spurious and due to the red dice being counted in both throws. Hence if we evaluate ρ it will give the correlation coefficient. Now, σ_x^2, σ_y^2, $\sigma_{y'}^2$ follow the Bernoulli frequency distributions and hence their values are tpq, $(s-t)pq$, $(s-t)pq$, respectively. Substituting these values in (3) we have

$$\rho = \frac{tpq}{\sqrt{(tpq + (s-t)pq)^2}} = \frac{t}{s},$$

which proves the theorem.

Theorem II. Given three sets of variables with a common subset,

[1] Yule, *Introduction to the Theory of Statistics*, 5th Ed., p. 211.

that is, $z_1=x+y$; $z_2=x+y'$; $z_3=x+y''$, then the partial correlation coefficients of the first order $r_{12,3}=r_{23,1}=r_{31,2}=\frac{t}{s+t}$ where t is the number in the common subset x and s is the total number in the sets z.

To prove this we need only note that since

$$r_{12,3}=\frac{r_{12}-r_{31}r_{23}}{\sqrt{1-r_{31}^2}\sqrt{1-r_{23}^2}}$$

and since the partial correlation coefficients of zero order $r_{12}=r_{23}=r_{31}=t/s$ then substituting their values in (4) we have at once $r_{12,3}=\frac{t}{s+t}$ and similarly for $r_{23,1}$ and $r_{31,2}$.

In the same way it is easy to prove the following two theorems.

Theorem III. Given $n+2$ sets of variables $z_1, z_2, \ldots z_{n+2}$, each having the common subset x, the partial correlation coefficients of nth order are equal to $\frac{t}{s+nt}$.

Theorem IV. Given three sets of variables with a subset common to each pair, that is, $z_1=x_2+x_3+x_4$; $z_2=x_1+x_3+x_5$; $z_3=x_1+x_2+x_6$; where t_i $(i=1, 2, 3)$ is the number in the common subset x_i, then the partial correlation coefficients of first order

$$r_{ij,k}=\frac{st_k-t_it_j}{\sqrt{(s^2-t_i^2)(s^2-t_j^2)}}. \qquad \begin{matrix}(i, j, k=1, 2, 3)\\ (i\neq j\neq k).\end{matrix}$$

Various kinds of urn schemata can be built on the three last theorems and carried out experimentally, similar to those of Darbishire which illustrate the first theorem. As an illustration of Theorem IV the following experiment in dice throwing has been performed here. Throw s dice of which t_1 are red, t_2 white, and t_3 black, $t_1+t_2+t_3=s$ and count the number of sixes that appear. Let the black dice lie, throw the other two colors and note the number of sixes again. For the third throw let the white dice lie, and throw black only. Then count the number of sixes that appear and also include the number of sixes that appeared on the red dice in the first throw. Then the partial correlation coefficients will be as stated in Theorem IV.

On finding the correlation coefficient between two sets of data, one must take precautions that no spurious correlation exists. For example, if one should correlate the marks made in a psychological test with the marks made on the arithmetic problems in the test, the result obtained will be too high. For if there were no correlation at all between them we would nevertheless obtain a spurious correlation,

amounting to the ratio of the number of arithmetic problems to the total number of problems in the test. This is due simply to the fact that the arithmetic problems are included in both sets of data. This is a point which the non-mathematical user of statistical methods is liable to overlook, and consequently he falls into error.

SEASONAL VARIATION OF DEPENDENCY

By F. Stuart Chapin,[1] *University of Minnesota*

The fluctuation of social phenomena with the business cycle has been the subject of many recent articles.[2] In 1918 Rubinow [3] computed the annual or seasonal fluctuation of fifteen series representing different forms of dependency in New York City for the calendar years 1914 to 1917. The data had been gathered by the newly organized Bureau of Social Statistics of the Department of Public Charities of New York City, and comprised complete material from certain representative relief agencies.[4] Although these data are not regarded by Rubinow as a complete description of poor relief phenomena, they appeared to be sufficiently typical of relief-giving agencies to represent fairly general trends in the fluctuation of demand for relief.

In general, Rubinow's dependency indexes (computed on the basis of the average monthly figures for 1916) fell into two different groups: (1) those with pronounced seasonal fluctuation showing considerable variation from year to year (secular trend) and extremely sensitive to economic crises; (2) those with seasonal fluctuation less marked showing different laws of seasonal variation and relatively constant from year to year.

Annual or seasonal fluctuations were computed from twelve of these series (for the years 1914–16) by Rubinow, who used a method of simple averaging. No allowance was made for the affect of secular trend upon the seasonal variation in the series that were complicated by a definite secular trend and a violent cyclical disturbance. It is the object of this paper to show the practical effects of recomputing the seasonal variation by allowing for the influence of the secular trend and

[1] Assisted by Charles R. Hoffer, Ernest J. Meili, and George A. Lundberg.

[2] References given by Ogburn, W. F., "The Fluctuations of Business as Social Forces," *Journal of Social Forces*, Vol. I, Number 2, January, 1922, pp. 73–78.

[3] Dependency Index for New York City, *American Economic Review*, Vol. 8, December, 1918, pp. 713–40.

[4] General family relief or case work of C. O. S., care of homeless men and women, operations of the municipal lodging house, number of free burials, committment of children to institutions by the Department of Public Charities, payment to private hospitals for dependent patients treated by them, amount of work performed by dispensaries, and number of small loans.

of the business cycle. If the corrected results show a different distribution of the annual load on the relief agency, then the practical value to the executive head of the agency of using refined statistical methods to ascertain the normal variation is established. We have approximated the secular trend for 48 months by three methods for each of three series: (1) the moving average; (2) the parabolic trend;[1] (3) the linear trend.

The table presents a comparison of Rubinow's figures for the annual variation of relief (free lodgings) with our figures, corrected by allowing for the influence of the secular trend and computed by a simple method of monthly means. This method was used in preference to Person's link-relative method because our data seemed to fit Hart's theorem.[2] It will be observed at once from an examination of the table that our figures differ considerably from Rubinow's. The difference is typified in the chart where Rubinow's annual cycle of dependency and his annual cycle for free lodgings are compared with our corrected seasonal variation for the latter data.

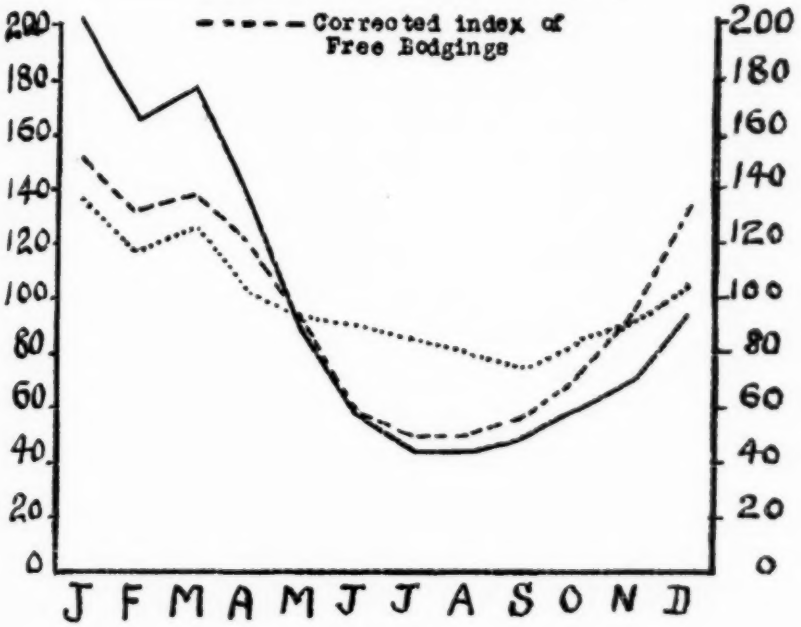
Before the practical significance of the effect of using a more refined statistical tool than Rubinow used is commented upon, there is a theoretical difficulty which deserves consideration. The use of the method of monthly means is predicated on the assumption that the series analyzed fulfils the following condition: Moderate cyclic influences well distributed along an approximately straight-line trend. In the case under consideration, the question may be raised whether the violence of cycles that reflect economic disturbances invalidates the use of the method of monthly means with its assumption of a linear trend. Should this criticism be valid we would expect to find a marked difference between the cycles computed from the line of least squares and those computed from the parabolic trend. An examination of the curves based on the deviation from 48-month trend lines for family relief, free lodgings, and homeless men and women shows that this is not the case; in fact, the curves on the linear and parabolic trends fit quite closely. We conclude, therefore, that the use of the method of monthly means with its linear assumption is justified in this case.

The practical significance of our corrected figures for the executive of a relief agency is shown in the table and particularly in the chart. In every case our corrected series shows an earlier incidence of the

[1] Family relief $Y = 231.42 - 5.6X + 0.06X^2$
Free lodgings $Y = 495 - 20.13 + 0.24X^2$
Homeless men and women $Y = 640.06 - 25.7X + 0.324X^2$
January, 1914, taken as point of origin. Parabolas centered.

[2] "The Method of Monthly Means for Determination of a Seasonal Variation," this JOURNAL, Vol. 18, N. S. 139, September, 1922, pp. 345, and Davies, G. R., "*Introduction to Economic Statistics*," pp. 116–20.

Seasonal variation in Dependency Index of New York City, seasonal variation in index of Free Lodgings, original and corrected for influence of secular trend and business cycle.



Months of the year. Typical seasonal variation based on 48 months from January 1914 to December 1917

COMPARISON OF SEASONAL VARIATION OF DEPENDENCY

	Family cases—number of new cases		Homeless men and women—number of applications		Municipal lodging house—total registration		Child commitment—number committed	Dispensaries—number of new patients	Small loans—number applied	Gen. index
	Rubinow	Corrected	Rubinow	Corrected	Rubinow	Corrected				
Jan.	146.9	113	163.2	131	202.8	152	102.0	103.6	95.8	138.8
Feb.	127.4	102	149.7	112	165.1	133	92.2	83.3	85.4	118.1
March	119.8	103	138.8	110	177.5	139	108.6	103.7	104.7	124.9
April	86.3	79	98.0	84	138.5	121	103.9	105.7	89.5	106.7
May	85.9	84	77.5	72	89.7	85	111.0	116.8	98.4	96.9
June	87.9	86	61.9	62	59.2	59	112.5	111.3	110.5	90.2
July	82.7	84	62.4	67	46.1	49	103.4	104.3	109.6	85.1
Aug.	73.9	80	63.0	71	44.1	49	95.8	104.8	98.3	80.7
Sept.	66.4	77	62.1	76	49.0	58	95.0	92.3	89.2	76.2
Oct.	81.9	101	80.9	106	60.0	74	97.1	99.3	101.1	85.1
Nov.	95.4	123	102.6	147	71.8	93	89.0	89.7	98.7	91.1
Dec.	144.7	191	117.7	187	96.3	132	89.7	85.2	109.4	106.1

NOTE: Columns 1,3, 5, 7, 8, 9, 10 are taken respectively from columns 1, 3, 5, 6, 7, 9, 10 of Rubinow's Table 4, page 737.

seasonal load. The effect of our correction has, therefore, been to push the peak load back into the winter months; that is, the high point of demand on relief agencies falls in November and December, rather than in December, January, and February as it would appear from Rubinow's series. This difference is to be attributed to the fact that our corrected seasonal cycle eliminates the influence of extreme business cycles from January, 1914, to May, 1915. The correction eliminates this abnormal tendency, while Rubinow's seasonal cycles computed on simple averages are unduly affected by the abnormal influence of these years of extreme disturbance.

To the executive who must plan to meet the peak load of the relief demand by providing adequate financial reserves and a larger staff, it is a matter of considerable practical importance to know that in the normal annual variation the peak load is likely to come in November and December rather than in January and February of the winter.

FRENCH-CANADIAN INTERMARRIAGE IN NEW ENGLAND

The ultimate test of the assimilation of an immigrant group lies in the extent of its intermarriage with other groups in the community. Such a test of the French-Canadians in New England is significant, for not only are they numerous and well established, but a definite effort is made by both church and national organizations to prevent racial mixture, which is called by one French publication "a crime against God and a national abomination."

The marriage records of five New England cities, Fitchburg, Leominster, Southbridge, Woonsocket, and Worcester, for the year 1921, were recently examined for the purpose of ascertaining the extent of intermarriage in this group. In each of these cities there is a large French-Canadian population of comparatively long residence. The cities vary in size from Worcester with 200,000 to Southbridge with about 15,000. In Worcester the French-Canadians and their descendants constitute less than a tenth of the population and stand fifth in the racial groupings; in Fitchburg and Leominster they constitute from one-third to one-fourth of the population and stand second among the racial groups, being exceeded by the English-American stock only; in Woonsocket, with approximately two-thirds of the population, and in Southbridge, with three-quarters, they are in the lead, with the American stock a poor second. Racial origin was determined by birthplace, birthplace of parents, Christian and family names, and occasionally by the nationality of the clergyman when other tests were doubtful. Some errors undoubtedly were made, but probably not many.

In all there were 843 marriages out of a total of 3,307 in the five cities in which one or both parties were French-Canadian by birth or extraction. Table I gives the percentage of intermarriage.

TABLE I

MARRIAGES IN WHICH BOTH PARTIES WERE FRENCH-CANADIANS COMPARED WITH INTERMARRIAGES IN THE FIVE CITIES IN 1921

Cities	Both French-Canadian		Intermarriages	
	Number	Per cent	Number	Per cent
Worcester	140	62.3	88	37.7
Leominster	40	70.2	17	29.8
Fitchburg	87	74.3	30	25.7
Woonsocket	287	79.1	70	20.1
Southbridge	76	80.7	17	19.3
Totals	621	73.7	222	26.3

The nationality of the groups with whom the 222 French-Canadians intermarried was also secured as follows: Americans (generally although not always English-Americans) 91, Irish and Irish-Americans 72, French (from France) 14, English-Canadians 9, Italians 8, Scandinavians 6, British 6, Belgian 5, Austrian 2, German 2, Poles 2, Finns 1, Greeks 1, and Russian-Jews 1. Of the 222 mixed marriages 49, or 22 per cent, were performed either by a justice of the peace or by a Protestant minister.

If the American group is omitted, in order that comparisons can be more readily made with the figures given by Dr. Drachsler for New York City,[1] the ratio of intermarriage among the French-Canadians is 18.9. This is not much larger than the ratio for all nationalities for New York City, 13.59. Since Drachsler's data include a large Jewish element, notoriously opposed to intermarriage, the resemblance is especially significant. Drachsler, however, found the percentage of intermarriage among the French-Canadians in New York to be 75.6. The striking difference is due unquestionably to the make-up of the populations studied; New England is a French-Canadian center, while the proportion of French-Canadians in New York City is infinitesimal. As compared with other older immigrant groups studied by Drachsler, for instance, the Germans with 53.05, or the French from France 49.55, the amount of amalgamation of the French-Canadians going on in New England is relatively small.

It would seem, therefore, that the mass immigration, the comparative proximity to the homeland, and the strong social pressure exerted have been reasonably successful in preventing as much amalgamation of the French-Canadians in New England as would naturally take place.

ROBERT C. DEXTER

PROGRESS OF WORK IN THE CENSUS BUREAU

CENSUS OF WEALTH, DEBT, AND TAXATION

The collection in the field of statistics for the wealth, debt, and taxation inquiry, the scope of which was fully described in the June issue of this JOURNAL, has now been completed for places to which it is necessary to send agents, such

[1] Drachsler, J., *Intermarriage in New York City*, Columbia University Studies, 1921.

as states, counties, and cities, and it is practically completed, through correspondence, for townships and other small organizations, such as drainage and irrigation districts. The report nearest to completion is that on taxation and revenue systems, which consists of a digest of the laws pertaining to revenues of the states and the local governments. This report is now in the hands of the printer.

Press summaries containing the principal statistics on revenues, assessed valuations, indebtedness, and tax levies are being published as fast as the figures can be tabulated. Figures have already been tabulated and published for one state, and a number of other states will be ready within a short time. There are approximately 65,000 political units with power to collect taxes or to incur debt. The facts for all these units are in the hands of the Bureau, and by the end of the calendar year the public will have the main statistics on these subjects.

Because of a variety of exemptions provided for by the laws of the different states, and because of the varying degrees in the completeness of the valuation of such articles as are not exempt, the assessors' figures of personal tangible property do not furnish a basis for an accurate statement of the value of such property. The Bureau of the Census, therefore, has to resort to other methods to discover the quantity of personal effects of various kinds in the several states and to estimate their value.

For live stock values the Bureau utilizes the estimates of the Department of Agriculture as to the value of live stock on farms. The value of farm implements and machinery is based upon the value of these articles as reported to the Census in 1920, with corrections for depreciation, new equipment, and changes in price levels. The value of manufacturing machinery, tools, and implements is based upon the latest census reports with corrections as indicated by examination of data for a number of the leading industries. The value of gold and silver coin and bullion is calculated by adding to the coinage value of the gold the market value of the metal in standard silver dollars and subsidiary coins, as given in the report of the Secretary of the Treasury. The value of stocks of agricultural, manufactured, and mining products are estimated from data compiled by the Department of Agriculture, the Bureau of the Census, and other governmental bureaus.

The estimate of the value of household and office equipment, clothing, and articles of personal adornment presents difficulties out of proportion to the values to be arrived at. To assist in securing a proper basis for this estimate a brief questionnaire (which the members of the Statistical Association have already received) was sent out in large numbers with a view to obtaining some data as to the average value of household furniture and wearing apparel per family in different sections of the country and for different classes.

CENSUS OF ELECTRICAL INDUSTRIES, 1922

The field work for these investigations was virtually brought to a close on June 30. There have been 58,775 schedules received to date for all branches of the electrical industries including 5,500 for electric light and power stations, 1,250 for electric railways, 52,000 for telephones, and 25 for telegraphs. These are being reviewed, errors are being corrected by further data received from the

companies, and the schedules being prepared for tabulation. A list comprising the names of 149 railway companies having telegraph or telephone lines has been compiled from the records of the Interstate Commerce Commission, and information in regard to the operation of these lines is now being secured by correspondence and will be included in the final report on telegraphs. In order to avoid making duplicate inquiries and to save time and expense, the schedules on file in the Interstate Commerce Commission covering the public utilities of telephones, telegraphs, and electric railways have been used as far as practicable in preparing the census schedules for those investigations. The additional information necessary to complete the census schedules and the reports of companies not making returns to the Interstate Commerce Commission have been taken up by correspondence.

BIENNIAL CENSUS OF MANUFACTURES, 1921

In addition to the 283 preliminary reports which were given to the press showing general statistics for each manufacturing industry, 57 bulletins will be issued presenting detailed statistics for the more important industries. Of these 19 have already been published and the manuscript or proof for 27 is now in the hands of the printer.

ANNUAL CENSUS OF FOREST PRODUCTS, 1922

The collection of data for the three annual reports on forest products—namely, lumber, lath, and shingles; pulpwood consumption and wood-pulp production; and turpentine and rosin—has been completed and the final tables and manuscript are being prepared. Each of these reports will be published in bulletin form.

EXTENSION OF ANNUAL AND SEMI-ANNUAL REPORTS

As a part of the plan of the Secretary of Commerce to furnish at more frequent intervals information pertaining to certain important commodities, and at the request of trade associations interested in the respective inquiries, the Bureau will hereafter compile and publish annual reports on the production and sale of farm equipment, clay and refractory products, and lighting fixtures; and semi-annual reports on sulphuric acid and acid phosphate used in the manufacture of fertilizer, and on the production of paint and varnish.

CENSUS OF MARRIAGE AND DIVORCE

The field work on the census of marriage and divorce covering the calendar year 1922 is also very nearly completed. For 10 states it was possible to secure the divorce reports from state records, but for the remainder of the country the data had to be secured through the aid of county officials or by Bureau agents. The divorce schedules, for the most part, were sent out early in the year, and by the end of June complete returns had been received from nearly 95 per cent of all the counties thus covered and for 7 of the 10 states for which data were furnished from state records. More than 140,000 divorce reports have thus far been received, and it is hoped that the entire number for the United States will

be available by the end of July. The report of marriages, consisting of a simple statement of the number of marriages performed during the year 1922, was obtained from state records for 25 states and from county officials for the remainder of the country. Marriage reports had been received by the end of June from more than 90 per cent of the county officials and for 23 of the 25 states in which the reports were secured from state records.

The last statistics of marriage and divorce collected by the Bureau of the Census relate to 1916, but collections had been made theretofore covering the 40 years from 1867 to 1906, those for the first 20-year period by the Department of Labor and those for the second 20-year period by the Bureau of the Census.

J. A. H.

DEPARTMENT STORE RESEARCH

DISCUSSION AT A DINNER MEETING OF THE AMERICAN STATISTICAL ASSOCIATION

A dinner meeting of the New York City members of the American Statistical Association was held at the Machinery Club on the evening of June 8. The meeting was under the direction of the Committee on Distribution and was devoted to the uses and problems of statistics in department stores. Two main papers were presented, one by Mr. Ernest Katz, Controller, R. H. Macy and Company, and the other by Mr. J. B. Swinney, Merchandise Manager of the Young Peoples Store, Lord and Taylor. The papers were discussed by Dr. Paul H. Nystrom, Director of the Retail Research Association, Mr. Michael Dowley, Merchandise Manager, Arnold, Constable and Company, Mr. Alvin Dodd, Director of the Retail Distribution Division of the United States Chamber of Commerce, and Dr. Oswald Knauth, Secretary of the National Bureau of Economic Research. About seventy-five members were present.

Mr. Katz discussed the topic from the angle of the controller of the largest cash store in the world, which does an annual business of approximately $55,000,000. Mr. Katz believes that the controller should have a body of well-trained statisticians whose business it is to gather intelligently such facts as will be useful to the executives of the store in determining store policies, and to merchandise managers. These statistics should indicate particularly trends and tendencies of a business. Mr. Katz thinks that many facts are now collected which have no bearing on business, and that certain collected data which have a bearing on business often are ineffective or useless because they are not interpreted. In the Macy store more than fifty types of statistical information are compiled. This information is given daily to an interpreter who in turn presents his conclusions to the controller for discussion with the operating committee of the store at its regular meetings. The reports of the controller are watched very closely by the executives of the store, and by merchandise managers and buyers, and are used as guides and indexes in determining store and merchandising programs.

The research study dealing with the delivery system brought out some

interesting facts. It has been generally conceded that the automobile is the most efficient and the cheapest means of delivering parcels. At Macy's store it was demonstrated after a series of studies that the horse-drawn vehicle was the most economical for nearby frequent stops; that the electric car is best for a second area just outside the first circle; and that the motor car excells only when used for distant points and infrequent stops.

Merchandise managers and buyers are restricted in help by the cost per $1,000 of sales. A certain cost is established, and department heads must adjust their departments to meet this cost.

Mr. Swinney was in entire accord with Mr. Katz in the latter's statement that merchandise managers should not maintain statistical departments. Mr. Swinney's experience in the retail field convinced him that the controller's reports are weather vanes indicating the size of stock and the volume of business during current weeks and should be compared with corresponding periods of other years. He believes that a careful weekly analysis of stocks and sales of all classifications of merchandise is essential.

The discussion of the papers brought out the fact that the large stores are making rapid progress in coördinating departmental work within the stores themselves and in coöperating with one another in the exchange of information.

We need a closer and more intensive study of the problems of distribution. The productive end of business has spent millions of dollars in learning how to produce goods efficiently and at the lowest possible cost. This country is suffering from a lack of scientific distribution. The wastes and the costs of distribution have raised the prices of some articles of merchandise to prohibitive levels, and they offer a fruitful field for investigation and study. The price of the article over the counter, whether it be food or other merchandise, represents too great a spread when compared with the price of the article at the factory or the price of the raw material on the farm. The discussion did not touch this phase of the problem. It was confined entirely to methods and principles involved in the work.

The speakers in my opinion presented the best exposition of the subject yet offered. The committee believes that this field will stand considerable exploitation and it will be glad to coöperate in the study of any topic bearing on the problem.

ANDREW H. MELVILLE, *Chairman*

MISCELLANEOUS NOTES

CHARLES FELTON PIDGIN

Charles Felton Pidgin, who was connected with the Massachusetts Bureau of Statistics of Labor for over thirty-four years, died at his home in Melrose Highlands, Massachusetts, on June 3, 1923.

Mr. Pidgin was born in Roxbury, Massachusetts, November 11, 1844. After graduation from the English High School, Boston, in 1863, he was employed as an accountant in mercantile business until his appointment, in 1873, as chief

clerk of the Massachusetts Bureau of Labor Statistics, in which capacity he served under Colonel Carroll D. Wright (1873–88) and Mr. Horace G. Wadlin (1888–1903). In 1903 Mr. Pidgin was appointed Chief of the Bureau to succeed Mr. Wadlin. He became the fourth Chief of the Bureau which was established in 1869.

Although from boyhood he suffered greatly as a result of an accident to his hip, and was partially blind for a number of years, he was an unremitting worker. In addition to performing his administrative duties as an official of the Bureau, he contributed much statistical material for publication in the Bureau's reports and labor bulletins, and in 1888 he published, independently, a book entitled *Practical Statistics*. He also invented numerous devices and machines for the mechanical tabulation of statistical data, some of which were patented. His spare time out of office was devoted to literary work, and he gained even greater prominence as an author than as a statistician. The most popular of the long list of his writings were his novel, *Quincy Adams Sawyer*, and his historical romance, *Blennerhassett*.

In 1907 Mr. Pidgin severed his connection with the Bureau of Labor Statistics, and up to a short time before his death he was engaged principally in literary work.

R. F. P.

A PROPOSAL FOR A CLEARING HOUSE OF INFORMATION ON THE EMPLOYMENT OF MEMBERS

Some educational institutions have maintained for their graduates a clearing house of information regarding positions open and positions desired. Any one who wishes may file with this bureau a statement of his qualifications, experience, positions held, salary desired, etc. It is not merely for those who wish to obtain positions, but for any one who especially desires to advance himself or make a change. It is often extremely difficult for a large corporation or other institution to find the kind of man that is especially required, and a rather laborious personal search has to be made to find any one who would care to make a change.

It has been suggested, and very cordially approved by the Board of Directors of the Association, that the Statistical Association inaugurate this kind of information service for its members by establishing a register of statisticians. The objects of such a register would be, first, to furnish prospective employers with the names of persons having experience in various lines of statistical work, and, second, to afford to the members of our Association a means of keeping in touch with the best opportunities for employment. It is desired at this time to learn the reaction of the members, favorable or unfavorable, to the proposal.

The plan involves keeping on file in the Secretary's office or in the hands of a special committee sufficient up-to-date information about each member desiring a position or a change of work to enable the office to determine whether the member would be interested in any opportunity which comes to our notice, and to determine whether an employer would care to look further into the qualifications of a candidate.

No responsibility would be assumed of vouching for the detailed qualifications of a man for a particular job, but the plan would serve to bring together available candidates and prospective employers.

At present no fee is contemplated and the Association would assume no obligation to secure positions, but every effort would be made to call the attention of employers to available candidates and their experience and general qualifications, *as represented by their own statements. The rest would be a matter of direct negotiation.*

Would you be interested in such a plan, and would you keep the office informed of the necessary details in case you use the service? *Without sending in your statements of desires and qualifications,* will you not express your opinions at this time before we start the actual service? Address R. E. Chaddock, Secretary, Kent Hall, Columbia University, New York City.

The International Institute of Statistics will resume its annual meetings. The next meeting will be held in Brussels during the first week in October. This will be the first meeting since the session at Vienna in 1913. The preparation of the program has been the work of a joint committee representing the International Institute of Statistics, the Economic Commission of the League of Nations, and the International Labor Bureau. The program will consist of the following papers and reports:

1. Statistics of International Commerce, by M. Armand Julin
2. Statistics of Mineral Production, by M. Methorst
3. Statistics of Agricultural Production, by Professor Rici
4. Statistics of Fisheries, by Sir Henry Rew
5. Index Numbers of Production, Prices, and Cost of Living, by M. Lucien March
6. Unification of the Statistics of the Causes of Death, by M. Huber
7. Unification of Infant Mortality Statistics
8. Progress of Vital Statistics in the United States since 1900, by Professor Walter F. Willcox
9. Individual Bulletins for Population Registration, by M. Methorst

The Statistical Council of the Bureau of Agricultural Economics of the United States Department of Agriculture was organized by Dr. H. C. Taylor, the Chief of the Bureau, to advise with him and with one another on statistical matters arising within the Bureau. Dr. Taylor felt that in an economic organization of such size there was need for uniformity in statistical methods with regard to research, analysis of data, and the form of presentation. Matters referred to the Council are assigned by the Chairman to sub-committees for intensive examination. The report of the sub-committee is then discussed by the Council at large, and a recommendation made to the Chief by the Chairman of the Council (Mr. Schoenfeld, Director of Marketing Research). As a result of the report of the sub-committee on machine tabulation, there has been established in the Bureau a machine tabulation service section combining the equipment of

the several divisions under the direction of Mr. Smith. The subject of index numbers has been under discussion since February. A sub-committee is now preparing a report on the method and procedure of expanding the index numbers at present issued by the Bureau. The members of the Council, all of whom are members of the American Statistical Association, are W. A. Schoenfeld, Chairman, O. C. Stine, G. C. Haas, W. F. Callander, S. A. Jones, R. H. Wilcox, S. W. Mendum, C. R. Chambers, B. B. Smith, H. R. Trolley, L. A. Adams, and J. A. Becker.

Work of the Statistics Committee, Public Safety Section of the National Safety Council. The Statistics Committee of the National Safety Council is preparing for the Safety Congress, which is to be held at Buffalo, New York, early in October (this year), an analysis of the available accident statistics for the United States for the years 1911–21. A special section will also be devoted to statistics of automobile fatalities in principal cities of the United States during 1922. Introductory to the report there will be an important section on the conduct of traffic movement surveys, the preparation of traffic accident maps and diagrams, and the graphic solution of some ten or twelve key problems in traffic control. The statistical section of the report will be prepared by Dr. L. I. Dublin, Chairman of the Committee, assisted by Mr. E. W. Kopf. The special automobile department will be developed by Mr. H. P. Stellwagen, and the traffic study section by Messrs. W. J. Cox and G. C. Kelcey. Another section will be devoted to the practical use of the standard accident reporting forms which were recommended for use by American cities in the Committee's report for 1922. Mr. Brownell will have charge of this part of the report. Mr. A. H. Kohler will prepare an important section on the practical results of uniform reporting as developed by locals of the National Safety Council. There is hardly any problem now being submitted to statistical analysis so important as this question of the practical prevention of fatalities on our public highways. It is hoped that the membership of the American Statistical Association will be sufficiently interested in this problem to offer suggestions on practical ways and means of studying and preventing traffic accidents in our cities.

Statistical Research in Cancer Mortality. The Statistical Bureau of the Metropolitan Life Insurance Company is conducting an intensive analysis of the cancer mortality data available in the records of the industrial department of their company. It is hoped that the results will be of unusual significance, in view of the growing interest in cancer control. The analysis will take into consideration the changes in recorded cancer mortality rates during the past decade, and comparisons will be made between the experience of the insured group and that for the population of the United States Registration Area, and perhaps for populations of certain foreign countries. The data are rich in detail as to color, sex, and age, and have been compiled for eleven consecutive years in accordance with uniform classification and tabulation procedure. The research is being conducted by Professor Ira V. Hiscock, of the Yale University School of Medicine, Department of Public Health, under the supervision of Dr. L. I. Dublin, Statistician for the Metropolitan Life Insurance Company.

The Division of Analysis and Research and the Office of Statistician of the Federal Reserve Board were consolidated on July 1, 1923, into one division under the title, Division of Research and Statistics. Mr. W. W. Stewart, who had been director of the Division of Analysis and Research since September 15, 1922, is the director of the new division.

An Advisory Statistical Council on Immigration has been formed by Mr. Cotillo, State Senator for New York and Chairman of the Joint Legislative Committee on Immigration. Miss Kate Claghorn is Chairman and Mr. Bradley Buell is Secretary of the Council.

At the June meeting of the Royal Statistical Society of London, Mr. W. S. Rossiter, President of this Association in 1922, was elected a Fellow. His election makes twenty-three American Fellows in the Royal Society (including honorary Fellows). The organization of the British Society in 1834 led to the formation of the American Statistical Association in the United States in 1835. It is thus due largely to the influence of the Royal Society that our own Association is the dean of practically all scientific organizations in this country. Of the total number of American members above mentioned, all but four are members of the American Statistical Association. The members of both associations are: Roger W. Babson, John B. Blood, Burdett Ross Buckingham, Arthur R. Crathorne, Otto R. Eichel, Irving Fisher, H. Pierson Hammond, Frederick L. Hoffman, John Koren, Edward B. Morris, Henry L. Rietz, William S. Rossiter, Frank R. Rutter, Harry A. Toulmin, Jr., Lee J. Wolfe, S. H. Wolfe, A. A. Young, Warren M. Persons, Walter F. Willcox.

Dr. W. H. Steiner, formerly Acting Chief of the Division of Analysis and Research of the Federal Reserve Board, has been appointed Assistant Professor at the College of the City of New York, in the School of Civic and Business Administration, where he will conduct courses in banking and finance. He will continue to serve as economist of the National Association of Mutual Savings Banks.

Mr. Burton Dunlop and Mr. Otto Richter, both graduates of the actuarial and statistical courses at the University of Michigan, have been appointed to the statistical staff of the American Telephone and Telegraph Company of New York City.

Dr. R. B. Robbins, formerly Assistant Professor of Mathematics at the University of Michigan, has returned to the Insurance Department of New York and will be stationed in New York City.

Professor Walter F. Willcox, of Cornell University, will attend the meeting of the International Institute of Statistics to be held in Brussels, during the first week in October, where he will read a paper on the "Progress of Vital Statistics in the United States since 1900."

Professor E. R. A. Seligman has returned from Europe where he has spent the past nine months studying the fiscal problems of some of the European countries. His work was in connection with the League of Nations.

Mr. W. L. Crum, formerly of the Department of Mathematics at Yale University, has been appointed Assistant Professor of Statistics in the Department of Economics at Harvard University, with part time assignment to the Committee on Economic Research.

Dr. Frank M. Phillips, who has been in Scientific Research in the United States Public Health Service for the past four years, has become Chief of the Division of Statistics in the United States Bureau of Education. Dr. Phillips is Consultant in Research for the Public Health Service and for the Bureau of Mines. He is also Professor of Statistics at George Washington University.

Dr. J. M. Chapman, instructor in banking at the Columbia University School of Business, has spent the summer in Washington doing special work for the Bureau of Foreign and Domestic Commerce on the *Commerce Yearbook*, a new publication issued by the Department of Commerce.

MEMBERS ELECTED SINCE MAY

Adams, Lawrence, Bureau of Agricultural Economics, Washington, D. C.
Anderson, Lawrence B., Guaranty Trust Co., New York City
Beardsley, Henry S., The Trenton Times, Trenton, Mo.
Brolyer, Cecil R., Stanford University, California
Brown, T. H., American Tel. & Tel. Co., New York City
Chambers, C. R., Bureau of Agricultural Economics, Washington, D. C.
Chinlund, Edwin F., Arthur Anderson & Co., Accountants, New York City
Coover, John Edgar, Associate Professor of Psychology, Stanford University, Calif.
Cox, George Clarke, Wood, Struthers & Co., Investment Securities, New York
Cunningham, John B., Michigan College of Mines, Houghton, Mich.
Dewhurst, J. Frederic, Wharton School, University of Pennsylvania, Philadelphia, Pa.
Everle, George J., University of Southern California, Los Angeles, Calif.
Fitzgerald, Jay E., U. S. Chamber of Commerce, Washington, D. C.
Fletcher, Brooks, Lecturer, Cleveland, Ohio
Foote, Frank N., Comptroller, Franklin Railway Supply Co., New York City
Gay, Constance M., Case Pomeroy & Co., New York City
Haas, G. C., Bureau of Agricultural Economics, Washington, D. C.
Handman, Max S., University of Texas, Austin, Tex.
Hering, Irwin M., L. F. Grammes & Sons, Allentown, Pa.
Killick, Victor W., Independent Petroleum Marketers Ass'n., Los Angeles, Calif.
Mendum, Samuel W., Bureau of Agricultural Economics, Washington, D. C.
Miner, Mrs. Zaida F., Teachers College, Columbia University, New York
Moorhouse, Harold W., Director of Economic Research, American Farm Bureau Federation, Chicago, Ill.
Newhall, Arthur B., Hood Rubber Co., Watertown, Mass.
Neifeld, Morris R., Psychological Studies, New York City

Patton, Alson C., Phoenix Mutual Life Ins. Co., Hartford, Conn.
Piper, Clarence B., School of Business Administration, Harvard University
Post, Arthur E., Federal Reserve Bank, Philadelphia, Pa.
Raubenbeimer, Albert S., Stanford University, California
Reeves, William H., School of Business, Columbia University
Rockwell, C. P., Great Southern Life Ins. Co., Houston, Tex.
Rosenthal, Morris S., Stein, Wall & Co., New York City
Sargent, Noel, Manager, Open Shop Dept., Nat'l Ass'n of Manufacturers, New York City
Seitz, Enoch B., Executive Sec'y, American Washing Machine Mfgs. Ass'n, Chicago, Ill.
Smith, Bradford B., Bureau of Agricultural Economics, Washington
Tolley, H. R., Bureau of Agricultural Economics, Washington
Vanderblue, Homer B., Com. on Economic Research, Harvard University
Weber, G. M., Institute of Economics, Washington, D. C.
Wilcox, R. H., Bureau of Agricultural Economics, Washington, D. C.

REVIEWS

The Biology of Death, by Raymond Pearl. Philadelphia: J. B. Lippincott Company. 1922. 275 pp.

The *Biology of Death*, by Dr. Raymond Pearl, Professor of Biometry and Vital Statistics in the Johns Hopkins School of Hygiene and Public Health, belongs to the field of biological philosophy rather than to vital statistics. Humanity is used by him to illustrate life, vital statistics being employed mathematically in ways common to biometricians. The book contains very few constructive ideas useful to those who are working to promote human health; in fact, health officers will find in it much to criticize, little to commend, and nothing to inspire. A book of this sort is not without its value, however, even to health officers, as an antidote for unjustifiable optimism. There are some bubbles which deserve to be pricked. For example, it is often said that the "span of human life" is being lengthened as a result of public health activities. This is a very misleading statement. It is a fact that the *average* duration of life is increasing, but this average includes all deaths, young and old. The rising *average* age at death is due to lower infant mortality and to a reduction of diseases common to early life. Persons who reach middle life today have no greater expectation of further life than they had many years ago. Is it fair to the man on the street, therefore, to tell him that the "span of life" is increasing? Would it not be truer to tell him that the span of life is not increasing, but that of the children born today a larger proportion grow to manhood and womanhood? Dr. Pearl does not say this in so many words, but he quotes interesting data showing changes in specific death rates and in the expectation of life which justify such a statement.

The main argument of the *Biology of Death* runs about as follows: Life, though a physico-chemical process, is inherently continuous. Unicellular organisms, which multiply, one cell into two, two into four, and so on, are potentially immortal, using this term in a biological sense—that is, they are immortal unless they are killed, as most of them are. With the appearance of multicellular organisms, involving a differentiation of functions among the cells, natural death became possible, for a functional breaking down of one organism causes the whole complicated organism to perish. Death is the penalty of specialization. The failures of particular organs to carry on are in accordance with physico-chemical laws and the factors which control them are inherited. Hence, inheritance is the great outstanding influence in longevity. Each organism has a specific longevity which can be modified within a limited range by changes in environmental conditions. The most important influence of environment on natural death is that of hastening or retarding the rate at which the inherited endowments are used up. The death of a highly developed organism, such as man, is an orderly biological process, not a haphazard affair. Efforts to improve the public health have an influence, of course, but they are not so great as many seem to think. A layman is inclined to ask why some of these statements need to be proved, while

a statistician, after carefully studying the author's data, may well doubt whether he has proved them.

Chapter III on the chances of death, life tables, and expectations of life, and Chapter VII, which describes the results of experiments on the duration of life of fruit flies, comparisons being made with the expectation of life of human beings, are the best parts of the book and very interesting.

In Chapter IV the causes of death are classified in a new way in order to show how and when the different organs of the body break down. In this classification bacterial, trophic, and mechanical causes of death are not separated. The organs which cause death are listed in the following order, according to data for the United States Registration Area, the figures showing their relative importance: Respiratory system (396), alimentary tract and associated organs (335), circulatory system and blood (21), nervous system and sense organs (176), kidneys and related excretory organs (107), primary and secondary sex organs (88), skeletal and muscular system (13), skin (10), endocrinal system (1.5), all other causes (171). Having made this new classification, the author gives the specific death rates by ages for each class for males and females. The results, plotted on semi-log paper, are interesting, but the attempt to draw general conclusions from such incongruous composites cannot be regarded as successful. Had the diseases been given separately, the data would have been much more useful. It should be remembered, however, that the author has his mind on the biological, not the public health, problem. The amount of work involved in making the composites must have been very great. A table is given showing the most fatal organ systems at different ages. For infants it is the intestinal tract; from age 1 to age 54 in males it is the respiratory system; after that the circulatory system; from age 1 to age 49 in females it is the respiratory tract; from 50 to 59, the alimentary tract, and after that the circulatory system. The point of the chapter is that each organ has a characteristic time curve for its breakdown. This is worth knowing.

In Chapter V the separation of diseases is pushed back into the embryo, the causes of death being grouped according to the endoderm, the mesoderm, and the ectoderm. Chapter VI is a study of inherited longevity.

In the last chapter the author compares death rates for four "controllable causes"—tuberculosis of the lungs, typhoid fever, diphtheria and croup, and dysentery—with four "causes not controlled"—bronchitis, paralysis without specified cause, purulent infection and septicemia, and softening of the brain—for a twenty-year period. Using semi-log paper, he plots the data and notes that the controlled diseases have decreased at no faster rate than the uncontrolled. Although he concludes from this only that "the trend of human mortality in time is an extraordinarily complex phenomenon," the comparison invites the unwarranted inference that control has not been effective. Medical men may well question a classification which implies that no direct attempt has been made to control bronchitis, purulent infection, and septicemia. Bronchitis, moreover, is not a definite disease, but a catch-all for various infections.

The last chapter contains a graphical analysis of population statistics. As a study in curve fittings it is interesting, but as a basis of estimating future popula-

tion its practical value may be questioned. No amount of mathematical theory can take the place of judgment as to the probable effect of future conditions, and future conditions are bound to be very different from those of the past and present. Dr. Pearl reaches the conclusion that the population of the United States will never exceed 200 million persons. He may be right, but his mathematical conclusion is worth no more and perhaps is worth less than that of the agricultural economist who claims that intensive agriculture coupled with changes in human diet may make it possible for 400 million people to live in the United States. The curve brings out one fact, however, which is of real and practical significance to the country, namely, that the rate of growth of the United States is falling off. We have passed the point of inflection of our population curve. Henceforth retardation must take the place of acceleration.

In his preface Dr. Pearl says his object was to stimulate thought. In this he has succeeded, and one may hope that some of the points brought out will be taken up for active discussion. Life extension is occurring at the younger ages. Is it well for the race to have the life of persons biologically unfit extended to the age when they can reproduce? What is to be the relative influence on population growth of the extension of the life of the young and of the postponement of marriage? Will one effect offset the other? If the altruism of the white race acts to extend the life of the young of the colored races, will this be accompanied by a reduction of their birth rate? If not, what will be the ratio of colored to white persons in the next century? Or again, if unicellular life is immortal and specialization brings natural death, what shall be said of our modern civilization based on specialization and super-urbanism as compared with the "unicellular" rural life of former days and still existing in certain parts of the world? Must our civilization die because of its specialization, and which of its many organs will break down first? The *Biology of Death* is certainly a suggestive subject for thoughtful study.

GEORGE C. WHIPPLE

The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability, by Irving Fisher. Boston: Houghton Mifflin Company. 1922. xxxi, 526 pp. (First publication of the Pollak Foundation for Economic Research.)

The progress of any science is dependent upon the establishment of standards and instruments of measurement. The lack of precise quantitative measures in the field of economics forms a serious obstacle to the verification of theory and the general advancement of the science. It seems particularly fitting, therefore, as the editor of the Pollak Publications observes in a prefatory note, that the first publication of the Pollak Foundation should be *The Making of Index Numbers*.

In this book Professor Fisher sets for himself two chief purposes: first, to investigate all the formulas for index numbers which have been or could reasonably be constructed in order to examine their relative reliability and determine the possible trustworthiness of a properly devised index number; and, second, to facilitate the computation of index numbers by selecting the best formulas and

recommending the simplest methods for their calculation. To these ends Professor Fisher develops an ingenious numbering system for the formulas he examines, classifying them according to the type of average and the method of weighting employed, and subjects 134 different index numbers to careful scrutiny.

As criteria for adjudging the relative merits of these formulas two fundamental tests are set forth, the time reversal test and the factor reversal test. The first of these, the time reversal test, requires that the ratio between one point of comparison and another point taken as a base shall be the reciprocal of the ratio obtained by the same formula when the given point and the base point are interchanged. Professor Fisher discards as theoretically unsound, however, the circular test (the requirement that fixed base and chain methods shall give identical results) by alleging that an index number properly compares two and only two points and that indirect comparisons by way of other points are invalid. According to Professor Fisher, the second necessary property of a good index number, conformity to the factor reversal test, is that the index numbers for two related factors shall give results consistent with the true movement of their product, as, for example, in the case of commodities the product of the index number of prices and the corresponding index number of quantities shall equal the true relative value.

It is found upon application of these tests that there is a bias, or tendency to err in a fixed direction, inherent in certain types of average and in certain methods of weighting. Formulas which possess opposing type and weight biases are termed erratic from their liability to err in either direction. Others are freakish because of their insensitiveness to changes in prices, quantities, or the number of commodities. The only formulas that conform to both tests are obtained by taking geometric means of certain simpler biased formulas or by averaging weights. Thus Professor Fisher is enabled to discard all but a few index numbers, and he further discriminates for minor considerations among those retained. From this group of satisfactory formulas, numbering 13 in all, one is selected on the score of accuracy as being the "ideal" index number. It may be described as a geometric mean of two ratios of aggregates, one with base year weighting, the other with given year weighting. In Professor Fisher's notation the "ideal" index numbers of prices and of quantities for the year "1" relative to the year "0" are, respectively,

$$P_{01}=\sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}\times\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \text{ and } Q_{01}=\sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0}\times\frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}}.$$

It is pointed out, however, that the two ratios of aggregates of which the "ideal" is the geometric mean are usually in close agreement, and it is recommended that the first of these aggregatives (that with fixed or base year weighting) be used in case the necessary weights for each year are unavailable.

This procedure, starting with very general premises, results in the selection of an ideal formula for all purposes. Consequently, the author's conclusion is that the purpose for which an index number is constructed does not affect the choice of formula: the index number that is best for one purpose is best for all purposes.

The book is designed to meet the needs of the beginner as well as those of the specialist in index numbers, and is so arranged as to permit of a hasty culling of the essential points without the necessity of careful and continuous reading. In view of the complexity of the materal treated, however, it would seem preferable to have presupposed some acquaintance with the fundamental principles of index numbers, and so to have eliminated much of the extremely elementary text which is likely to annoy the more advanced class of reader. As a concrete illustration of the characteristics of each of the 134 formulas, the index numbers of both prices and quantities based on 1913 are computed for a group of 36 commodities for each of the years 1914 to 1918, inclusive. The immense amount of labor involved in this computation will be duly appreciated by those who have had experience in the preparation of index numbers. Finally, the author discusses certain practical considerations, including the representativeness of samples and the effect of errors in the data, and gives a table showing the relative speed of calculation of each formula.

Professor Fisher's conclusion that the choice of formula is independent of the purpose for which an index number is constructed is in reality a condition to the validity of his two reversal tests, for his principal justification for each of these tests can be undermined by denying that the purpose is immaterial. As a basis for the time reversal test he asserts that "no reason can be assigned for choosing to reckon in one direction which does not apply to the opposite"; and for the factor reversal test, "no reason can be given for employing a given formula for one of the two factors which does not apply to the other." As a useful instrument an index number implies a purpose, and for different purposes the requirements are dissimilar. To one investigator, for example, the temporary importance of a commodity as measured by the total value entering into trade at a given time may be of prime interest, whereas another investigator with different motives may desire to construct an entire series of index numbers with a fixed set of weights based on values during a period which he considers normal. Some makers of index numbers will not concede that for their purposes momentary changes in quantities should influence an index number of prices.

Many users of index numbers are primarily interested in the *movement* of a curve, that is, in comparisons of each point not only with the base but also with every point in the series. Here the circular test must be fulfilled, although in a slightly different sense from that in which Professor Fisher interprets it. For he insists on consideration of quantities in each of the years under examination. If, however, one wishes to employ a constant weighting throughout the series of price index numbers, regarding changes in quantity as irrelevant to the problem, the circular test can be fulfilled in the sense that comparisons of changes between all points are rendered significant.

The matter in which we differ fundamentally with Professor Fisher, then, is that of weighting. He maintains that whenever there is a price of anything exchanged there is implied a quantity of it exchanged; so that the problem of an index number of prices implies the twin problem of an index number of quantities. Usually, however, the maker of index numbers is concerned with only one of these measurements, and in selecting weights he endeavors to represent the

importance of the individual items during some normal period of time, not necessarily identical even with the base period. The manufacturer who wishes to follow the composite trend of the prices of the materials in which he is interested, for the purpose of planning future purchases, may adopt weights based on the quantities he normally purchases over a long period of time. An index number which would be affected by month-to-month fluctuations in his actual purchases might be seriously misleading; indeed, it is conceivable that an index number which would conform to Professor Fisher's two tests might show a startling change while the price of every single commodity in the index remained the same. For certain purposes a simple average of price relatives may be appropriate, as in the case of the stock market speculator who wishes to obtain a picture of the fluctuations in the prices of a selected list of stocks. To him each stock presents an equal opportunity for profit. The number of shares outstanding, the size of the floating supply, and the volume of trading in each stock have no direct bearing upon his problem of measurement.

These instances, we hold, are proper fields for the application of index numbers. According to Professor Fisher, an index number may be defined simply as an average of relatives; but an index number is "good" only in so far as it conforms to his two reversal tests. In this requirement there is implied a purpose, which assumes that two factors are involved and that the measurement of either is dependent in part on the other. The definition of an index number and the restrictions placed upon a "good" index number are inconsistent in that index numbers which live up to these tests may be good for some purposes but are not good for all purposes.

As an extensive and orderly compendium of information about index numbers this book forms an important and a much-needed addition to the literature of the subject. Professor Fisher is highly to be commended for undertaking so exhaustive and painstaking a study, and the Pollak Foundation for making this great work possible. The book will well repay careful examination by all students of index numbers.

DONALD R. BELCHER and HAROLD M. FLINN

Public Relief of Sickness, by Gerald Morgan. New York: The Macmillan Company. 1922. 195 pp.

Mr. Morgan's purpose is to show that in a system of state health insurance, such as was proposed in New York and other states in 1920, the high cost of the medical benefits necessarily results in efforts at economy which seriously impair the value of such service. Since a health system cannot furnish adequate service, the solution lies in creating an independent state-controlled health service in the form of "health centers." By using the results of the statistical studies of the Illinois Health Insurance Commission of 1919, Mr. Morgan first shows the great importance of medical service to the groups usually covered by compulsory health insurance systems. He then reviews the experience of these systems in Denmark, Germany, and Great Britain, and finds that in each of these countries the medical services leave much to be desired. As an alternative he proposes

that in any American system it would be desirable to arrange for a complete divorcement of the cash benefits from the medical benefits. The insurance organization, he contends, should be restricted to furnishing cash benefits; the medical benefits, consisting of the physician's services, hospital treatment, the supply of medicines, appliances, etc., should be provided by a state system of health centers, entirely independent of the insurance organization (though co-operating with it) and available for the entire population.

This plan of course has much in its favor—in fact, one is inclined to wonder why the author did not confine himself to making a plea for the inauguration of such a state medical service without combining it with a discussion of compulsory health insurance. His argument in favor of the state "health center" system is an excellent presentation of the reasons for its adoption, while his discussion of the actual workings of health insurance in foreign countries contains many statements which, to say the least, indicate a tendency to discredit them. Thus his description of the British medical service (on pages 87 to 93) is made up largely of quotations from Brend (four out of seven pages) in spite of the conspicuous bias of this writer. No reference whatever is made to the reports of the Chief Medical Officer of the Ministry of Health, in which that officer gives an account of the administration of the medical benefits of the system and incidentally shows what an erroneous impression is given by Brend's exaggerations. Also, if these reports had been consulted, Mr. Morgan would have found an account of a series of cases of anthrax caused by imported shaving brushes (Annual Report of Chief Medical Officer, 1919–20, pages 263–267) and he would hardly have made the statement (on page 137) that contracting anthrax from a shaving brush at home "would be so very unusual as to be considered impossible."

This dependence on secondary sources of information also leads the author to make the doubtful statement (page 49) that "the Danish death rate at last reports (1921) was about 19 per 1,000, which is exceedingly high"; the latest available Danish statistical year book, that for 1922 (on page 16), gives only the 1920 rate, which was 12.9 per 1,000—a very favorable rate. Such errors as these seriously impair the value of this argument for a plan which has much in its favor.

HENRY J. HARRIS

Social Change With Respect to Culture and Original Nature, by William F. Ogburn. New York: B. W. Huebsch, Inc. 1922. viii, 365 pp.

In spite of a slow-moving and somewhat cumbrous style, Professor Ogburn has written one of the most suggestive books of the year in the field of social theory. Written in the finest spirit of scientific candor, the book reveals in its author an objectivity toward his own theorizing that might well serve as a model for more confident but less sagacious social philosophers. The volume is consistent in viewpoint (though the arrangement is not severely logical) and hence its contents may be reviewed under a few main propositions. The viewpoint is that of the school of cultural determinists which has greatly increased in vogue latterly.

The argument begins with the assertion that the role of heredity, or of the psycho-biological factors, in social evolution has been exaggerated. It then contends that during the past 25,000 years in which there have been immense cultural changes, human nature has remained substantially the same. Consequently, cultural advances are not to be explained by any improvement in human qualities on the average, but by factors operating uniquely within the historico-cultural environment itself. The author finds that "theoretically, cultural changes may be accounted for in terms of changes in the biological nature of man or in purely cultural factors" (page 61). Cultural changes tend to be cumulative and hence continuous advance can be made without any improvement in the human capacity to invent or to originate. It is true that inventors come from the upper ranges in the distribution of mental abilities; moreover, "an improvement in the inherent mental ability of the race would result in an increased number of inventions"; but no such improvement is demonstrable. Hence, since the bio-psychic basis of inventions is certain to remain pretty constant, and since "the existing culture is the mother of inventions," it is sound sociological methodology to account for "social evolution in terms of cultural records before recourse is had to the more obscure biological causes." (See especially pages 62, 63, 73 *et seq.*, 80, 81, 83, and 130–142.)

In view of the epoch-making work of certain American anthropologists in this field no one would be disposed to question the immense value of the point of view here presented. Moreover, the author's frequent use of "may" and "might," evidencing a consciousness of the complexities of the matter, disarms criticism. But, while admitting that we have here an important field of inquiry and a very suggestive point of approach, inevitable differences of emphasis arise.

In the first place, the question whether human nature is the same today, yesterday, and always is in many respects the most important question in social science. An argument based on a few paleolithic skulls is certainly not adequate on this point. It should not escape attention that the average of inherited qualities of two groups may be the same in spite of striking *differences in the range and the proportions of the statistical distribution of their qualities about the average*. When the author says that "there were very probably changes in human nature preceding the last ice age, but it has not been proven that there have been any changes since," one is tempted to reply, "nor has it been proven that there have not been." Moreover, the average size of the human skull might be the same at two widely different dates in spite of numerous and great *oscillations* in both average size and statistical proportions at every point of the range of distribution between these dates. The human inheritance of the Greeks may well have been better than that of many of their neighbors or of modern Mexico, even though the average capacity of their skulls (of which we know all too little) was less than that of the Cro-Magnons (of which we must also speak with slight assurance) or no greater than that of modern Englishmen. In other words, there may be "biological changes" without "biological evolution" or mutations.

I think we must feel a little vagueness also in the author's use of the term human nature. This he uses in a general way to describe the inherent qualities.

He leaves one to assume that he means to include all man's organic powers as well as his intellectual, emotional, and character traits. He finds that Europeans have grown in height and weight during the last few centuries, "but this may be due to purely environmental influences" (122). Here arise a variety of questions: Is it possible to maintain with any confidence that all the different powers and traits of man have been about the same on the average in all times and places? Even if so, can one also maintain that the varied *combinations* of these qualities, in all the kaleidoscopic amalgamations of racial elements, have also been on the average the same? If man has increased in height and weight within a few centuries, is it reasonable to suppose that he has not changed also in some or all of his psychic powers? If man must be viewed as a domestic animal, would there not be in the course of 25,000 years a sort of up and down movement in his qualities in correlation with the extent of his domestication? Would not any such changes due to domestication act on the evolution of culture in the same way as inherited mutations? Finally, it should be kept firmly in mind that all men are human; that it is, therefore, possible to say that human nature is the same among all humans wherever and whenever found; but that the real question is whether the statistical distribution of these human traits when measured either singly or in combination shows not merely the same average and the same range of distribution but the same proportionate frequency at all points in the range.

As regards invention and cultural changes the important question is the relative frequency of those extreme variates who constitute the truly great men. A number of such are enumerated on page 90 *et seq.* among those making simultaneous discoveries. Such men do not occur at a greater frequency than a few to the million of those born. It is a purely gratuitous assumption that every human group will be equally fecund in such men. In fact, one can say with a good deal of confidence that the evidence runs the other way. The fact that in a number of instances two or more persons make similar discoveries at the same time shows that culture suggests inventions, that culture grows, when it grows at all, in a cumulative fashion. But it does not prove anything as to the inevitableness of the matter. The necessary genius to profit by cultural suggestions may not have appeared. This the author admits when he finds that the rate of cultural evolution may be faster or slower, depending on the biological factor. Nor does history lend any support to the doctrine that great men will arise when cultural crises require them. One may add also that the outstanding geniuses tend to give what Ward called a sympodial character to cultural evolution.

The theory of the cultural determinists amounts to the assertion that cultural evolution is independent of its hereditary or racial background. But this is certainly questionable. Woodworth's well-known and oft-cited studies, while indicating some differences between racial groups in sensory powers, threw almost no light on racial differences with respect to the higher mental processes. Psychological measurements indicate that it is just with respect to these processes that the greatest differences occur between different grades of ability. The work of the mental-testers uniformly indicates significant differences in the distribution of mental capacities, as between social classes and racial groups.

Unless one is to assume dogmatically that the average qualities of all human groups is the same, he must be prepared to recognize that war, migration, conquest, differential birth and death rates, isolation and environmental selection, and mutations may cause changes in the content and potentialities of group germ-plasms. For example, contrary to the race purists there is much evidence of both historical and biological sorts for the contention, frequently made by Professor Giddings, that racial heterogeneity is an important condition for the development of a high civilization. There is no ground for supposing that the phenomenon of heterosis, so frequent in crosses in both plant and animal worlds, should be wholly wanting in human hybridism (see East and Jones, *Inbreeding and Outbreeding*, N. Y., 1919). On the other hand the crosses of morons of different races would not be productive of genius: it is not hybridism alone that matters; the quality of the crosses is also important. For this reason he would be highly gifted with temerity who would assert that no changes are now taking place in the inherent qualities of the American people, in the range and proportions of distribution. Even the race equalitarians can no longer assert that all levels of the population are equally fecund in superior types of individuals. It is because of these probable shifts in the qualities of a population that the author's comparison of the rate of biological evolution with that of cultural evolution fails to convince. Mutational changes in man may have come nearly to an end; so may they have also in cattle, but one would not conclude that therefore two herds of milch cows were equally endowed with desirable traits.

When Professor Ogburn admits that an improvement in inherited mental ability would increase the rate of invention, he implies that a decline in ability would check the flow of cultural changes. It seems to follow that if this decline went far enough, cultural changes would cease altogether. This points to a somewhat different conception of cultural dynamics from that which the author in general upholds. His position in the main is that there is something dynamic in culture itself, that it flows on and on because of some indwelling spirit, uniquely and peculiarly cultural. One cannot have his cake and eat it too, here any more than elsewhere. The fact would seem to be that when the human factor is withdrawn the dynamics of cultural evolution goes with it. One may not question the fact that the existing stage of culture is an important, an essential, condition in the stream of inventions; but it is only suggestive; it has no active creative power in itself; it comes to naught in something over 950,000 out of every million minds that come in contact with it. Is it not the neglect of this fact that has prevented the author from advancing any explanation for the fact that culture is not persistently cumulative but has led a very checkered career of ups and downs during the long period under survey?

This also raises some doubt as to the correctness of the author's conclusions regarding sociological methodology. He would study cultural changes in terms of cultural antecedents and omit speculations regarding that mysterious element, human nature. By all means let us try every possible method. But it should not be overlooked that bio-metric and psycho-metric methods and results are the most exact and tangible we yet have. Culture is a distinctly human product; no other animal produces it; it is limited at every point by the capacities of the

human animal, and derives its character at all times and places from the traits of its creator. Moreover, the author (especially page 336 *et seq.*) suggests the impossibility of controlling cultural evolution or even of predicting the social effects of any invention. We fully agree, and think of prohibition as an illustration. But we find in this a strong reason for centering the attention in social policy on the quality of the human stock by means of quality tests for immigration and reproduction, segregation and sterilization of the unfit, birth control propaganda at public expense, etc. These are tangible and can be made effective. They may not enable one to predict the future of civilization, but they constitute the surest means for its preservation and continued elevation. If this be so, then sound sociological methodology will not neglect them.

This review is already so long that space remains only to say that there are many other interesting propositions in this suggestive book. If exception is taken to any of them it would be by way of supplementing them from the viewpoint indicated in the foregoing paragraphs. Professor Ogburn has greatly enriched our sociological literature by putting in so stimulating a form the theory of the school of cultural determinists.

F. H. Hankins

Smith College

Report on the Problem of Raw Materials and Foodstuffs, by Corrado Gini, with Annexes prepared under his direction [by Professor Vinci and Dr. Sloutski]. Geneva: Printed by Albert Kundig for the League of Nations. [1922?] 251 pp.

This report, dated August 16, 1921, originated in the following resolution adopted by the Council of the League of Nations on October 25, 1920, following a report of the Italian delegate, Signor Tittoni:

> The Council has fully considered the difficulties experienced by numerous countries in assuring the import of raw materials essential to their welfare and even to their existence, and has requested the Economic Section of the Economic and Financial Committee to study:
>
> (a) The extent and nature of their requirements;
>
> (b) The causes of these difficulties (other than those arising from lack of credit or fluctuation in the rate of exchange, which have already been considered by the Brussels Financial Conference); the effects of the existence of monopolies will be very specially considered.
>
> The Council invites the Committee to submit to it at the earliest possible moment a report on the results of its enquiry, a report which is indispensable for the further deliberations of the International Economic and Financial Conference.

In the following January the Committee sent out to the various governments two questionnaires, one dealing with import and export restrictions, monopolies, and prohibitions, the other asking for statistical information with regard to the supplies of raw materials and foodstuffs. At the request of the Secretary General, Professor Gini undertook to devote six months "to drawing up a report on the production and distribution of raw materials and foodstuffs" to be laid before the Committee. This report was based not merely upon the results of the first questionnaires but upon the results of another more detailed one which

he sent out, upon personal visits which he made to fifteen countries of Europe, and upon other statistical data which he and his collaborators were able to gather. Professor Gini undertook not so much to comply with the exact terms of the Council's resolution as to give "the personal opinions of an expert on the whole question of the supply of raw materials and foodstuffs" and "to furnish the Committee with information which may be of use to it in its future labors."

The body of the report, written by Professor Gini himself, touches briefly upon the economic situation in the first half of 1921 and summarizes briefly the causes of European difficulties in obtaining supplies. Finding the situation strikingly different from that which prevailed when the inquiry was first proposed in 1919–20, he is not content to discuss the emergency but treats rather the perennial problem of which he considers this a passing phase. Most of his report he therefore devotes to "the possible steps which might be taken by the League of Nations."

In this discussion he goes rather far afield and assumes the role of political adviser to the League. At almost every point he urges the extension of the League's activities, even to the point of taking action upon national restrictions on immigration. Even friends of the League would regard many of his suggestions as impractical and unwise, and there are few upon which the League has acted.

The annexes to the report contain the bulk of the material dealing with the problem of raw materials and foodstuffs. These consist of brief summaries of the "position" before, during, and since the war with respect to seven groups of foodstuffs and raw materials, namely, cereals, wool, cotton, coal, mineral oil, iron industry, and chemical manures. Some fifteen to thirty folio pages are given to each group. Data on world production, trade, and consumption, with special reference to leading sources of supply and leading importing nations, are drawn largely from recognized sources and apparently only in slight degree from the replies to the questionnaires. The material is generally taken at its face value and is discussed simply and clearly. Considering their scope and brevity, the reports are well done, and they afford a convenient introduction to the subject for persons not already familiar with it.

Neither in the annexes nor in Professor Gini's report proper can one find the treatment (which the Council requested) of the extent and nature of the current requirements of the various nations and their difficulties in filling these requirements. Obviously the questionnaires, which were intended to yield specific information on these points, were ill-advised, unduly formidable, and well-nigh useless. Even under stress of war, when many lines were largely under government control, it was exceedingly difficult to obtain prompt and reliable statistical evidence of national supplies and requirements. In peace such a task is infinitely more difficult. Lacking such evidence, Professor Gini reported in sum: "I cannot tell you what you ask, but here is something that may be of interest on the same general subject."

JOSEPH S. DAVIS

The Development of Federal Reserve Policy, by Harold L. Reed. Boston: Houghton Mifflin Company. 1922. viii, 352 pp.

Mr. Reed has written an interesting and clearly-developed volume about the Federal Reserve system, and it will serve a useful purpose in spreading a general knowledge of our banking mechanism among readers who otherwise might not acquire more than a rather hazy idea of it. The concept of the work is particularly good, it being, as the title indicates, a discussion of reserve system policies rather than a mere description of machinery. If the technical execution of the book were as good as its plan, the outcome might easily have been something noteworthy. Even as it stands, the work is likely to be of value to various classes of readers.

It is not, however, very easy to see just why Mr. Reed arranged the volume in its actual sequence of chapters. He begins naturally enough with "The Regional System." But this chapter is followed by "Check Collections and Clearances," and the latter by "State Bank Membership." Next comes a treatment of "Rediscounts," and then a chapter on "Collateral Advances." After a chapter on the "Trade Acceptance," a section on "Agricultural Credit" is sandwiched in, and this is followed by a chapter on the "Bank Acceptance." A more curiously illogical topsy-turvy arrangement or sequence of chapters could scarcely be imagined; and much the same quality persists throughout the entire book.

But it is not the arrangement of material that will chiefly interest the reader of the volume, though that is of such a nature as greatly to lessen its worth as a "second book to follow the reading of one of the elementary texts in college courses"—a secondary purpose that was present in the author's mind. The book is in some degree of service for such an end, but its utility in that connection is likely to be limited. Its true use is to instruct the general reading public which wishes to know more of the reserve system, and for that end the monograph has a value that is not much affected by the technique of arrangement. For this class of readers the need is rather that of a series of brief clear papers or essays on the various phases of the system's operation. These essays need not have a very close integral connection; indeed, the "t. b. m." often resents such connection and prefers a book that gives him his data in broken doses.

What is the content of Mr. Reed's volume? It is fundamentally a running outline in rather sketchy form of some of the chief matters dealt with by the Federal Reserve Board in the development of its policy. Collections, acceptances, open market operations, deposits and reserves, and a variety of other elements in the Federal Reserve plan of banking are briefly outlined in successive chapters. The author usually begins with a short statement of "historical background" or financial theory, or both, follows it with a statement of the provisions of law applicable to the matter in hand, and then discusses with brief statistical commentary the chief events in the development of the policy relating to the topic under discussion. This is unquestionably the right method of approach for the more general reader, and if the content of the chapters were as good as their "lay out" the book might be an unusual production. As it stands, the writer has given in nearly every case a readable and fairly accurate account

of the matter he is treating. The chief fault lies in an uncritical attitude—not due to any effort to fawn or to flatter but to the hasty acceptation of statements apologia and efforts to make the worse appear the better reason as if all were equally trustworthy.

Added to this uncritical point of view as a serious defect in the book must be cited some hasty judgments here and there which the author will doubtless revise in later issues. "While progress in popularizing the trade acceptance was slow, it had proceeded upon a sound basis. . . . To secure the speediest mobilization of our war resources the funds should have been obtained from the existing volume of money and credit. . . . The general policy of the Federal Reserve administration during the period of our participation in the World War may be epitomized briefly as one of complete and cordial coöperation with the Treasury. . . . Had the physical volume of production been less large in 1919 and 1920 the date of the reaction might have been postponed." These are examples of the somewhat broad and thoughtless generalizations with which Mr. Reed's pages are punctuated. They show an immaturity of judgment and in some cases an inconsistency which will no doubt in due time remedy itself but which for the moment constitutes something of a blemish upon an otherwise useful restatement of Federal Reserve experience.

The more attention and discussion that is given to the Federal Reserve system the more likely we shall be to reach some early and definite adjustment of our banking problems. The questions involved have become complex and difficult to a degree that prevents the average man from following them clearly. Whatever can be done to stimulate public interest and to clarify thought on the whole issue should be done, and would be a real public service. Mr. Reed has collaborated in this important task and is entitled to due credit for the contribution he has made to thought on the subject.

H. Parker Willis

Selling Expenses and Their Control, by Northwestern University School of Commerce Bureau of Business Research, Horace Secrist, Director. New York: Prentice-Hall, Inc. 1922. 416 pp.

This is the final volume of an exhaustive study made by Professor Secrist during the last few years of retail clothing distribution under varying conditions. The book contains a wealth of material for the retailer, the student, the credit man, and the consumer. It was made possible by the coöperation of the National Association of Retail Clothiers in getting members to fill out questionnaires, and is another shining example of the value of trade associations. Similar studies have been made for such lines as shoes and jewelry by the Harvard Bureau, although no attempt was made to cover such details as this study has.

Previous studies by Professor Secrist showed that it was necessary to classify stores by years, by size, and by location, and this classification is therefore followed in the latest volume. The study was undertaken for two primary purposes: first, to determine what it costs to sell clothing in stores operating as independent units, and, second, to find, if possible, the principles governing these

costs and the methods by which they may be controlled. The results of the study are stated by the author as showing that "the relations of selling expense to sales and to total operating expense are governed by conditions which, to a large degree, are subject to effective merchandising control; that they vary according to clearly defined principles; and that these principles can be stated and the conditions of expense distribution under them measured."

The book is very timely in that clothing costs are rising and numerous charges of profiteering are hurled at the retailer. This study sets certain standards—rather rough, to be sure, but they will suffice for most practical purposes. In compiling his standards Professor Secrist might be criticized for not making greater use of frequency tables instead of showing simple averages. It would indeed be very interesting to know what the most efficient merchants' standards were for the various ratios. There are merchants whose knowledge of accounting is confined to the fact that merchandise should be marked up 50 per cent, while on the other hand there are others using the most modern methods of cost control. This is especially true of chain clothing stores, which are ignored in this study.

The analysis of selling expense and its components is based upon the records of 154 stores for the years 1919, 1918, and 1914. In 1914 selling expense constituted $11.48 per $100 total net sales; in 1918, $10.94; and in 1919, $10.44. In 1914, $7.52 of the $11.48 was spent for wages and salaries, $2.61 for advertising, and $1.35 for other selling expenses. In 1918 the amounts were $7.37, $2.22, and $1.35. In 1919 they were respectively $6.98, $2.22, and $1.24. For the combined years the total selling expense is larger for the stores in the larger cities (those with a population of 40,000 and over) than for those in the smaller cities.

The years chosen for the study have a very determining effect on the results inasmuch as two of the years were very abnormal years, 1914 being the year of the outbreak of the war and its demoralizing effect on business, and 1919 a year of post-war inflation. We would naturally expect selling expense to be somewhat lower during 1919, whereas during a poor year such as 1914 merchants advertised mark-down sales extensively and selling expenses mounted. Inasmuch as there are no "normal" years in business, the best that can be done is to take an average of these good and bad years in sales.

Inasmuch as there is no uniform accounting system in use in clothing stores, it is very questionable whether the primary sources of most of these data are reliable. One merchant might consider his window decorator's salary as a wage item and another as advertising, and so on down through the items. It was very interesting to note that in stores where the merchant used the commission basis in determining wages, the wage item was materially reduced per $100 of sales. Other studies, such as the Harvard Bureau's, cover profits, and it is unfortunate that they could not have been covered in this study. However, from the balance sheets that are available, it is known that the retailer is now operating on a surprisingly small margin.

Much complaint is heard today of the high rents that many clothing stores are paying, and the question is raised as to whether these stores have to mark up

their merchandise accordingly. This study shows that as the rent per square foot of floor space increases, selling expense and wages and salaries decrease. One result of increasing rental payments is found to be a saving in the amount of floor space utilized.

In making a study of selling expense in clothing stores it is a very serious question whether stores should be classified simply according to volume of business or according to the class of goods sold. Of course it would be very difficult to group stores in this second way, as many stores sell many grades of merchandise. To compare the selling expenses of a store selling cheap machine-made clothing with some of our exclusive New York men's stores doing the same volume of business would be very misleading. And in practically all cities clothing stores, as a rule, single out a class of trade which they desire and confine themselves to that class. In these high grade stores the work of the salesmen is very important, and perhaps the wage item here would differ materially from that of a cheaper store where price sells the merchandise.

With 292 tables the reader is swamped in a wealth of detail made necessary by the numerous operating conditions and classes of clothing stores. The average retailer, not statistically inclined, would find these tables too formidable for his use. This is unfortunate as they contain much of practical value. However, for class-room purposes and for the up-to-date retailer the book is indeed well worth reading.

W. J. Baxter

Clothing Forecaster Corporation

Total Permanent Disability Benefits in Relation to Life Insurance, by Arthur Hunter. Actuarial Studies, No. 5. New York: Actuarial Society of America. 1920. vi, 58 pp.

This study is one of a series being issued under the auspices of the Actuarial Society of America for the purpose of supplementing standard texts like the *Textbook of the Institute of Actuaries* and thus making more readily available to students and actuaries contributions on actuarial subjects widely scattered in periodical literature. The main purpose of the series is to assist students in actuarial science.

This study of total permanent disability benefits in life insurance was begun several years ago and in fact completed, in its first draft, before the European war began. The revision which is now offered to actuarial students and to the profession gains greatly through the great development of this type of insurance in more recent years. In 1910 there were, to be sure, many life insurance companies issuing total permanent disability clauses, but at that date the business was new and the experience of the companies with the clause was limited. The intervening years, up to the present, have brought about an enrichment in the quality of these clauses that has resulted in their becoming, no doubt, a permanent and valuable adjunct to life insurance protection.

No more appropriate selection could have been made for the author of this study. In addition to his high standing as an actuary, Mr. Hunter was one of

the three or four men who did pioneer work in formulating the actuarial basis of this type of insurance. He was assisted in the work by Mr. Mervyn Davis.

For the student and actuary to whom the book is addressed primarily, the parts of main interest will be the construction of combined mortality and disability tables and the derivation of formulas for determining premiums and reserves for the various types of benefits. The reviewer offers no judgment on these portions but depends on the professional standing of the authors and the sponsorship of the Actuarial Society as a guarantee of the quality of the work. For the non-technical reader the earlier, descriptive portions of the book, comprising the first 16 pages, will be of greatest interest. It is found here that although varying in detail, three main types of benefits are granted: (1) waiver of future premiums only; (2) waiver of premiums and payment of the sum issued in equal installments over a specified period of years, the balance on insurance payable in one sum if death occurs before the completion of the installments; and (3) waiver of premiums and the payment of an annuity during disability or to the maturity of the contract, and the payment of the face value of the policy at maturity or death. Of (2) and (3) the latter is obviously the more satisfactory, inasmuch as it furnishes a disability income and leaves the life insurance benefit intact. There are cases of course in which (2) will furnish all the benefit that is desired; but in general it seems that the (3) benefit is the one which most nearly meets the economic situation of a totally and permanently disabled person.

Some interesting figures are given, bearing out the conclusions of earlier piecemeal evidence, to the effect that about 70 per cent of all cases of total and permanent disability are due to tuberculosis, insanity, paresis, and paralysis.

BRUCE D. MUDGETT

Tables of Applied Mathematics in Finance, Insurance, Statistics, by James W. Glover. Ann Arbor, Michigan: George Wahr. 1923. xiii, 676 pp.

These tables are designed primarily for use in connection with college textbooks on the mathematics of finance, insurance, and statistics. They are admirably adapted to this purpose. Furthermore, the tables are of such wide scope as to be of great use to the student after graduation; they are sure to be of practical service to any one who has occasion to make computations in compound interest, insurance, or statistics, particularly to bankers, engineers, actuaries, and statisticians.

The cost of the book may at first thought seem to be high to some students, but the price is very reasonable when one considers the cost of production of such tables and the practical value of the book apart from its use in college work.

The material of the book is organized into four parts. Part I gives the value of compound interest functions to eight places of decimals for thirty-two rates of interest. The logarithms of the values are given to seven places. So far as the reviewer has been able to learn, these are the only published tables giving the logarithms of interest functions on such a large scale. Moreover, the rates of interest are selected with a view to practical application. For example, the values at interest rates of 7/12 per cent and $\frac{2}{3}$ per cent are included particularly for building and loan association computations.

The conversion tables for relating interest payable in p installments per year to the effective rate per annum are the most extensive tables which have been published.

Of special interest is the twenty-place table of values of logarithms $\log_{10}(1+i)$. This table is of importance in the construction of fundamental tables.

Part II gives the values of life insurance and disability functions. The life tables for a single life include commutation columns, valuation columns, life annuities, forborne life annuities due, temporary life annuities, single premiums for temporary and endowment insurances. The joint life tables include Hunter's Makemized American Experience Tables with commutation columns, life annuities due, single premiums and annual premiums for two equal ages at an interest rate of 3½ per cent. A valuable feature of the life tables is the life functions at 5 per cent for the United States Life Tables for white males and white females in the Original Registration States for 1910. These auxiliary tables will tend to acquaint the student with the applications of a table based on the general mortality rates of the country. The inclusion of Hunter's Combined Table of Mortality and Disability with corresponding commutation symbols will make the disability functions much more available than they have been and will facilitate elementary calculations in connection with the total and permanent disability benefit now very commonly offered in connection with life insurance policies.

Part III gives probability and statistical functions. This part will be very useful to students of the mathematics of statistics in applications to biology, economics, psychology, education, agriculture, and other sciences. One of the most important of these tables is that giving for the normal probability curve not only the area and ordinate but also the derivatives from second order to the eighth order. This table will particularly facilitate the applications of the Charlier frequency curves. The table of logarithms of the Gamma Function serves a similar purpose in the application of Pearson's generalized frequency curves.

The processes of interpolation and summation by the use of formulas involving finite differences will be greatly facilitated by the use of the values of the coefficients which have been included in pages 412–30. The tables give also the squares, cubes, square roots, cube roots, and reciprocals of integers from 1 to 1,000.

A unique feature of this book is the inclusion of an extensive table for the determination of the parameters of the exponential growth curve $y=ar^x$ corresponding to n observed statistical ordinates.

Part IV contains a seven-place table of logarithms of numbers from 1 to 100,000. Such a table of logarithms is practically indispensable to students of mathematics of finance. Logarithms to less than seven decimal places for the purpose of interest calculations are usually unsatisfactory when the results are to be carried out to the nearest cent, and often when they are to be carried only to the nearest dollar.

In conclusion it should be said that the student of mathematics of finance,

insurance, and statistics will find that this book of tables fills a real need, and any one who has to make accurate numerical calculations in any field of science will almost surely find these tables of service in facilitating the calculations.

H. L. Rietz

University of Iowa

RECENT LITERATURE

STATISTICAL BOOKS

Bergengren, Roy F. Coöperative banking. A credit union book. New York: Macmillan. 1923. 400 pp.

Bliss, James H. Financial and operating ratios in management. New York: The Ronald Press. 1923. xii, 396 pp.

Chapman, John M. Fiscal functions of the Federal Reserve Banks. New York: The Ronald Press. 1923. ix, 213 pp.

Comish, Newel H. The standard of living. Elements of consumption. New York: The Macmillan Company. 1923. xiv, 340 pp.

Edie, Lionel C., Editor. The stabilization of business. New York: The Macmillan Company. 1923. xii, 400 pp.

Foster, W. T., and Catchings, W. Money. Boston: Houghton Mifflin Co. 1923. viii, 409 pp.

Hamilton, Walter, and May, Stacey. The control of wages. New York: The George H. Doran Company. 1923. 185 pp.

Harvard University. Graduate School of Business Administration. Operating expenses in the wholesale grocery business in 1922. Cambridge: Harvard University Press. 1923. 47 pp. Bulletin No. 34.

Kelley, Truman L. Statistical method. New York: Macmillan. 1923. xi, 390 pp.

Klein, Philip. The burden of unemployment. A study of unemployment relief measures in fifteen American cities, 1921–22. New York: Russell Sage Foundation, 1923. 260 pp.

Middleton, Thomas H. Food production in war. New York: Oxford University Press. American Branch. 1923. xix, 373 pp.

National Bureau of Economic Research. Business cycles and unemployment. Report and recommendations of a Committee of the President's Conference on Unemployment including an investigation made under the auspices of the National Bureau of Economic Research. New York: McGraw-Hill Book Co. 1923. xl, 405 pp.

Redmayne, R. A. S. The British coal-mining industry during the war. New York: Oxford University Press. American Branch. 1923. xv, 348 pp.

STATISTICAL PUBLICATIONS

Canada. Dominion Bureau of Statistics. Prices and price indexes, 1918–22. Ottawa. 1923.

League of Nations. Economic and Financial Commission. Report on double taxation submitted to the Financial Committee by Professors Bruins, Einaudi, Seligman, and Sir Josiah Stamp. 1923.

Massachusetts. Department of Labor and Industries. Annual report for the year ending November 30, 1922.

New York Charter Commission. Report . . . with a draft of charter for the City of New York. 1923.

United States Government.

— Bureau of Education. Bulletins, 1923.

No. 10. The public school system of Arkansas. Part 1. Digest of general report.

United States Government.

——Bureau of Education. Bulletins, 1923.

No. 14. Public education in Oklahoma.

No. 18. Medical education, 1920–22, by N. P. Colwell.

—— Bureau of Labor Statistics. Bulletins.

No. 312. National health insurance in Great Britain, 1911 to 1921, by H. J. Harris.

No. 324. Wages and hours of labor in the boot and shoe industry: 1907 to 1922.

No. 325. Union scale of wages and hours of labor, May 15, 1922.

No. 327. Wages and hours of labor in woolen and worsted goods manufacturing, 1922.

No. 329. Wages and hours of labor in the men's clothing industry, 1911 to 1922.

—— Children's Bureau. Publications.

No. 93 (Revised edition). Child labor. Outlines for study. 1923.

No. 117. The welfare of children in bituminous coal mining communities in West Virginia, by N. P. McGill. 1923.

No. 118. Standards of public aid to children in their own homes, by F. Nesbitt. 1923.

No. 120. Maternity and infant care in a mountain county in Georgia, by G. Steele. 1923.

—— Public Health Reports. Volume 38. 1923.

No. 17. The prevention of simple goiter. Extent of rural health service in the United States.

No. 18. Studies on oxidation-reduction. Interesting population changes in the United States, 1910–20.

No. 22. Physical care of rural school children.

No. 23. Indices of nutrition.

No. 24. Provisional birth and mortality rates for 1922.

No. 26. Studies on the permeability of living and dead cells. Incidence of venereal diseases among American seamen in the Orient.

—— Women's Bureau. Bulletins.

No. 30. The share of wage-earning women in family support.

No. 31. What industry means to women workers, by Mary Van Kleeck.

ARTICLES IN PERIODICALS

American Journal of Sociology. 1923.

July. L. L. Bernard: Invention and social progress. J. H. Kolb: A study in rural community organization. A. W. Small: Some contributions to the history of sociology.

American Labor Legislation Review. 1923.

June. C. B. Walker: The twelve-hour shift. J. B. Andrews: Progress of American labor legislation. E. E. Witte: Treble compensation for injured children. T. I. Parkinson: Minimum wage and the constitution. F. E. Hering: Awakening interest in old age protection.

Giornale degli Economisti e Rivista di Statistica. 1923.

April. F. Zugaro: Le egemonie militari in Europa avanti e dopo la guerra. V. Alfieri: Le rilevazioni amministrative. G. de Francisci Gerbino: Di una proposta di riforma del diritto successorio.

May. M. Fanno: Inflazione monetaria e corso dei cambi.

June. V. Dentamaro: Il commercio estero dell'Italia durante la guerra. G. Majorana: Della dinamica patrimoniale nella presente economia capitalista.

Giornale di Matematica Finanziaria. 1923.

April. E. Blaschke: Il problema dell'interpolazione nell' assicurazione vita. P. Smolenski: Le teorie della riserva matematica nell'assicurazione vita.

International Labor Review. 1923.

April. S. Hansson: The trade union movement in Sweden. R. Sand: Health and the workers. Migration movements throughout the world in 1913, 1920, and 1921. F. Rager: Apprentice training in the Austrian metal industry.

May. Odon Por: The guild movement in Italy. C. Dulot: The present position in the French trade union movement. Adaptation of wages to cost of living in Poland.

Journal of Accountancy. 1923.

May. E. Fraser: To obtain sinking-fund payments from compound-interest tables. B. Rose: Responsibility of auditors. T. N. Beckman: Sources and treatment of surplus. W. T. Sunley: Minority interests in inter-company profits.

June. American Institute of Accountants' Examinations, May, 1923. M. E. Peloubet: Mechanical difficulties in consolidating accounts. O. B. Thayer: Receipt of certificates of capital stock as taxable income.

July. G. C. Carson: Elimination of inter-company profits in consolidated statements. G. H. Johnson: Amortization and depreciation in public service corporations. F. Vierling: The Franklin trust.

Journal of the Institute of Actuaries. 1923.

March. Opening address by the President, W. P. Phelps. W. Palin Elderton: Notes on the treatment of extra risk. W. Palin Elderton and H. J. P. Oakley: Report on the results of an investigation of the mortality experience of life annuitants during the period 1900–20.

Journal of the Institute of Bankers. 1923.

May. H. Hart: Gilbart lectures, 1923, II and III. J. Brunton: Some comparative notes on bankruptcy and winding up.

June. Annual report, 1923. H. Hart: Gilbart lectures, 1923, IV.

Metron. 1923.

Vol. 2. No. 3. A. A. Tschuprow: On the mathematical expectation of the moments of frequency distributions in the case of correlated observations. A. L. Bowley: The precision of measurements estimated from samples. W. Winkler: Die Berechnung der Tauglichkeitsprozente in der Heeresergänzungsstatistik. A. MacDonald: A study of the United States Senate. M. Boldrini e A. Crosara: Sull'azione selettiva della guerra fra gli studenti universitari italiani. A. De-Pietri Tonelli: Le fonti internazionali della statistica commerciale: II. America.

Monthly Labor Review. 1923.

May. A. L. Whitney: Health and welfare activities in the Government Printing Office. L. E. Tilden: New England Textile strike.

June. G. M. Price: Administration of labor laws and factory inspection in certain European countries. Compulsory labor service in Bulgaria.

Le Musée Social. 1923.

April. L. Bernard: Les Etats-Unis d'Amerique et l'Hygiène.

May. A. Isaac: La natalité française.

June. R. Merlin: Jules Siegfried. Sa vie. Son oeuvre

OTHER BOOKS

Demangeon, Albert. L'Empire britannique. Étude de géographie coloniale. Paris, Librairie Amand Colin, 1923. viii, 280 pp.

MISCELLANEOUS REPRINTS AND PAMPHLETS

Anthracite accounting and finance. Statement by the Anthracite Representatives of the United Mine Workers of America. Washington, 1923.

Armour-Morris purchase. How the savings and economies incident to enlarged business will benefit producers and consumers of meat. Published by Armour and Co., Chicago.

Calcul des probabilités. Sur le problème du schema des urnes. Note de A. Guldberg.

Commercial rent as an expense and its relation to profits, by Horace Secrist. Northwestern University, Bureau of Business Research, 1923.

Contributions to the analytical theory of sampling, by S. D. Wicksell. Stockholm, 1923.

Evidence in the case for the open shop. National Association of Manufacturers, New York, 1923.

Farmers' coöperation in Minnesota, 1917-22, by H. B. Price. Bulletin 202, University of Minnesota Agricultural Experiment Station, 1923.

Maps and diagrams, showing present conditions in New York and its environs, March, 1923. Prepared by the Physical Survey, Plan of New York and its environs, 1923.

Money in the United States. Bulletin No. 1 of the National Monetary Association, New York.

Organization and management of local livestock shipping associations in Minnesota, by E. W. Gaumnitz and J. D. Black. Bulletin 201, University of Minnesota Agricultural Experiment Station.

Pigmentation in the old Americans with notes on graying and loss of hair, by A. Hrdlicka. From the Smithsonian Report for 1921, pp. 443–484.

The possibility of extending human life, by L. I. Dublin. Metropolitan Life Insurance Co., 1922.

The Providence zone plan. Report outlining a tentative zone plan for Providence, by Robert Whitten, Consultant, to the Joint Standing Committee on Ordinances. Providence, 1923.

Recent improvement in the negroes' mortality, by L. I. Dublin. Reprinted from *Opportunity*, Vol. 1, No. 4, April, 1923.

The Rockefeller Foundation. A review for 1922, a summary for the first decade, by G. E. Vincent. New York, 1923.

The science of man: its needs and its prospects, by Karl Pearson. From the Smithsonian Report for 1921, pp. 423–41.

Testimony of Professor Irving Fisher: Reprinted from Proceedings of the Select Standing Committee on Banking and Commerce of the House of Commons of Canada on Bill No. 83, an act respecting banks and banking, and on the resolution of Mr. Irvine, M. P., re Basis, function and control of financial credit, etc. Ottawa, 1923.